National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

University of Alberta

Conformational Studies of Peptide Fragments from Decorin/DS-PGII by ^{1}H NMR, CD and Molecular Modeling

by

Yunjun Wang ©

A thesis
submitted to the Faculty of Graduate Studies and Research
in Partial fulfillment of the requirements for the degree of
Master of Science

Department of Chemistry

Edmonton, Alberta

Spring, 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-10762-0

Canada

University of Alberta

Library Release Form

Name of Author: Yunjun Wang

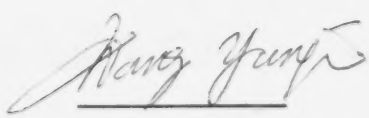
Title of Thesis: Conformational Studies of Peptide Fragments From Decorin/DS-PGII by ^{1}H NMR, CD and Molecular Modeling

Degree: Master of Science

Year this Degree Granted: 1996

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Yunjun Wang
Chemistry Department
University of Alberta
Canada, T6G 2G2

Date: Jan 3 1996

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled " Conformational Studies of Peptide Fragments From Decorin/DS-PGII by ^{1}H NMR, CD and Molecular Modeling " submitted by Yunjun Wang in partial fulfillment of the requirements for the degree of Master of Science.

George Kotovych

Dr. George Kotovych

Dr. Hsing-Jang Liu

Dr. R. E. D. McClung

Dr. Paul G. Scott

Date: Jan 3, 1996

Abstract

The conformation of four peptides (N-terminal acetylated and unacetylated 14-mers DEASGIGPEEHFPENH$_2$ and 24-mers AcQKGLFDFMLEDEASGIGPEEHFPENH$_2$ with a normal and an oxidized methionine residue), containing the sequence Asp-Glu-Ala-Ser-Gly-Ile-Gly, which is known to play an important role in the xylosylation reactions involving the glycosaminoglycan-bearing serine of decorin/DS-PGII, were studied by two-dimensional NMR techniques, circular dichroism spectroscopy and molecular dynamics calculation in a 60:40 mixture of methanol and water. The 14-residue peptide comprises the first 14 amino acids of the mature decorin protein and the 24-residue peptide incorporates an additional (N-terminal) sequence of 10 amino acids derived from the procore of decorin. The resonance heterogeneity induced by the existence of *cis* prolines (Pro^8, Pro^{13} in the 14-mer, and Pro^{18}, Pro^{23} in the 24-mer) in the studied peptides was evaluated from TOCSY and NOESY NMR spectra. The *trans-trans, trans-cis and cis-trans* isomers exist in approximate 68:25:7 proportions in methanol/water mixture. The NOE distance constraints were used as input parameters for restrained molecular dynamics and energy minimization calculations. It was demonstrated that the conformation of the DEASGIG fragment was affected by the presence of the ten amino acids at its N-terminal. The results indicate that an α-helix present in the 24-mer, beginning at the N-terminal with Lys^2 of the 24-mer and ending about Gly^{15}, and suggest that this could be the signal for the xylosylation of serine. A type VIb β-turn was observed involving the C-terminal *cis*-proline in the sequence His-Phe-Pro-Glu.

ACKNOWLEDGEMENTS

I would like to acknowledge some of the individuals who have helped me with this research over the past three years. First, I would like to thank my supervisor, **Dr. George Kotovych**, who has given me the opprotunity to carry out research in his laboratory, and who has given me intellectual direction and emotional support over the years.

I would like thank **Dr. Paul G. Scott** for his advice, his efforts in peptide synthesis and purification, and for his interest in collaborating on the projects.

I would like to give my appreciation to **Dr. Tom Nakashima** and other members of the NMR lab for their technical assistance.

I would like to thank **Dr. Jan Sejbal** for his help, advice and cooperation in carring out the molecular dynamics calculations.

Financial support by the **Department of Chemistry**, University of Alberta is gratefully acknowledged.

Table of Contents

Chapter	Page
1. Introduction	1
Structures in small peptides	1
Spectroscopic methods used to detect secondary structures in small peptides	3
Molecular simulation	6
References	10
2. Conformational studies of peptide fragments from decorin/DS-PGII by ^{1}H NMR, circular dichroism and molecular modeling	16
Background of this project	16
Materials and methods	18
Results	22
Discussion	28
Conclusions	30
Proposal for futher work	31
References	33

List of Tables

Table		Page
Table 1.1	Dihedral angles of hydrogen bonded β and γ turns	12
Table 2.1	Proton chemical shifts and NH temperature coefficients of peptide D14	36
Table 2.2	Proton chemical shifts and NH temperature coefficients of peptide AD14	37
Table 2.3	Proton chemical shifts and NH temperature coefficients of peptide OQ24	38
Table 2.4	Proton chemical shifts and NH temperature coefficients of peptide NQ24	39
Table 2.5	Constraints used for the TC isomer of peptide AD14 for molecular modeling	40
Table 2.6	Constraints used for the TT isomer of peptide OQ24 for molecular modeling	41
Table 2.7	Backbone torsion angles based on the minimized averaged structures obtained from molecular dynamics	42

List of Figures

Figure		Page
Figure 1.1	Standard nomenclature for the atoms and the torsion angles along a polypeptide chain	13
Figure 1.2	Pulse sequence for some 2D NMR experiments used in this paper	14
Figure 1.3	Schematic diagram of the steps used in this thesis for peptide structure determination by NMR	15
Figure 2.1	Glycosylated sequences in some small DS-PG	43
Figure 2.2	The structures of the 14- mer peptide, AD14	44
Figure 2.3	The structures of the 24-mer peptide, NQ24	45
Figure 2.4	Circular dichroism spectra of peptides D14 and NQ24	46
Figure 2.5	The αH-βH region of the NOESY spectra of peptide AD14	47
Figure 2.6	The αH-NH and βH-NH regions of the TOCSY spectra of peptide AD14	48
Figure 2.7	The NH-NH regions of the TOCSY spectra and βH-NH regions of the NOESY spectra of peptide AD14	49
Figure 2.8	Schematic diagram of *trans, cis* proline and the assignments of its protons	50
Figure 2.9	Sections of the NOESY and TOCSY spectra of the *cis* Proline[13] of peptide AD14	51

Figure 2.10 The αH-NH region of the NOESY spectra of peptide AD14 52

Figure 2.11 Net structural shifts of the α-protons for peptides (a) AD14 and (b) D14..... 53

Figure 2.12 The αH-βH region of the NOESY spectra of peptide OQ24 54

Figure 2.13 The NH-NH region of the NOESY spectra of peptide OQ24..... 55

Figure 2.14 Sections of the NOESY spectra of peptide OQ24 56

Figure 2.15 Net structural shifts of the α-protons for peptides (a) OQ24 and (b) NQ24..... 57

Figure 2.16 Survey of NOE effects recorded for the peptides (a) AD14, TT isomer; (b) part of AD14,TC isomer; (c) OQ24, TT isomer 58

Figure 2.17 The last 20 conformers from the 100 ps dynamic run for peptide AD14 with *cis* proline in position 13..... 59

Figure 2.18 Averaged minimized structure for the 24-mer with all bonds trans modeled without solvent..... 60

Figure 2.19 The stereo view of the average minimized structure of the 24- mer after modeling with the shell of solvent 61

List of Abbreviations and Symbols

CD	circular dichroism
FID	free induction decay
HPLC	high performance liquid chromatograph
MC	Monte Carlo
MD	molecular dynamics
mM	milli-molar
nm	nanometer
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enchancement
NOESY	two-dimensional nuclear Overhauser enchancement spectroscopy
PEF	potential energy function
ppm	parts per million
rf	radiofrequency
TOCSY	two-dimensional total correlation spectroscopy
UDP	uridine diphosphate

Naturally Occurring Amino Acids

A, Ala	Alanine
C, Cys	Cysteine
D, Asp	Aspartic acid

E, Glu	Glutamic acid
F, Phe	Phenylalanine
G, Gly	Glycine
H, His	Histidine
I, Ile	Isoleucine
K, Lys	Lysine
L, Leu	Leucine
M, Met	Methionine
N, Asn	Asparagine
P, Pro	Proline
Q, Gln	Glutamine
R, Arg	Arginine
S, Ser	Serine
T, Thr	Threonine
V, Val	Valine
W, Trp	Tryptophan
Y, Tyr	Tyrosine

Chapter 1. Introduction

Structures of Small peptides

Many peptides have been found to adopt a few regular secondary structures that are also found in natural proteins. The discovery that peptide fragments of proteins in solution, especially in water, do, in many cases, exhibit significant conformational preferences for recognizable structures(1-4) was of great interest and significance in several fields, notably those that sought an insight into the earliest events that initiate protein folding(5-7) and into the likely mechanism of induction of antipeptide antibodies (8).

Because of the great flexibility, the dihedral angles for the side chains are difficult to define. So secondary structures of a peptide usually are referred to as regular arrangements of the backbone of the polypeptide chain and can be described by a set of dihedral angles, ϕ,Ψ and ω (9), as indicated in Figure 1.1. The peptide bond has partial (40%) double-bond character, ω is very close to 180°, and C^{α}_{i} and C^{α}_{i+1} are all *trans* to each other. *Cis* peptide bonds, with ω=0°, can occur perhaps 25% of the time before proline residues but essentially rarely before any other residues. Hence, ϕ and ψ are the essential parameters required to describe the secondary structure of the peptides (10,11).

Secondary structures of peptides can be classified into three classes in order of their frequency of appearance. These are : 1. Reverse turns and nascent helices, 2. Helices, 3. β-sheets.

Reverse turns are the simplest regular conformation formed by linear peptides and can be classified into the β-turn and γ-turn, i.e., turns of four or three residues, respectively. These turns may or may not be stabilized by an intra-turn hydrogen bond; in β-turns this occurs between the C=O of residue i and the NH of residue i+3, whereas in γ-turns it is between the C=O of residue i and the NH of residue i+2. β-turns can be further classified as types I, II, III, their mirror images I′, II′, III′ , VIa and Vib both of VIa and Vib turn have *cis* proline at i+2). γ-turns can be further classified as regular and inverse γ turns. The classifications of β and γ turns according to their ϕ, Ψ angles are listed in Table 1.1.

The term nascent helix has been coined to describe an ensemble of turn-like structures over several adjacent residues of a peptide. Nascent helical structures typically occur in peptide fragments derived from helical segments of proteins (2).

The α-helix has 3.6 residues per turn, with a hydrogen bond between the C=O of residue i and the NH of residue i+4. The backbone conformational angles for right-handed α-helices are approximately ϕ= -60°, ψ=-60°, which is a favorable local conformation with a steep energy minimum. The only other principle helical species besides the α-helix which occur to any great extent is the 3_{10}-helix, with a three-residue repeat and a hydrogen bond to residue i+3 instead of i+4 and its backbone conformational angles are approximately ϕ= -60°, ψ= -30°.

A β-sheet is made up of almost fully extended strands which can interact in either parallel or antiparallel orientation, and each of the two forms has a distinctive pattern of hydrogen bonding. The antiparallel sheet has hydrogen bonds perpendicular to the strands and narrowly spaced bond pairs alternate with widely spaced pairs. A parallel sheet has evenly spaced hydrogen bonds which angle across between the strands.

Spectroscopic Methods to Detect Secondary Structures in Small Peptides

Several spectroscopic techniques(12) namely, CD, Raman, FTIR, and NMR , have been utilized for elucidating the solution conformations of linear peptides. Different spectroscopic techniques have different time-scales and the various forms of optical spectroscopy, for example, can distinguish species that interconvert on a subpicosecond timescale, whereas in NMR spectroscopy exchanging species are distinguishable only if they interconvert on a millisecond to a second timescale. Only *cis-trans* isomerization of the peptide bond, usually at proline, is slow enough for two distinct species to be observed in the NMR spectrum. NMR parameters reflect a population-weighted average over all conformations; populations can only be deduced indirectly, if at all. In this thesis, CD and ^{1}H NMR were used.

Two-dimensional NMR has been used in recent years as a structural tool for proteins and peptides (13). All 2D NMR pulse sequences may be described in terms of a number of time periods (14), as shown in Figure 1.2A. These are : a preparation period t_0; an evolution period t_1; a mixing period τ_m and a detection period t_2. The 2D NMR experiment also includes a mixing period or a mixing pulse. The preparation period usually consists of a delay time, during which thermal equilibrium or a steady state is attained, followed by one or more radio frequency pulses that create the desired coherence. During the evolution period the coherence of the system is frequency-labeled in a manner which depends on the length of the evolution period. The mixing period may include one or more rf pulses and delay intervals. During the mixing period, τ_m, information is transferred between spins by some physical process, so that connections between nuclei labeled in the evolution period and

nuclei detected during the detection period are established. During the detection period the system evolves further and the resulting free induction decay (FID) is recorded. The two fundamental types of 2D NMR experiment rely on different mixing mechanisms: coherent transfer requires the existence of through-bond spin-spin coupling, whereas incoherent transfer requires mutual relaxation from through-space dipolar interaction or chemical exchange of the spins.

Total correlation spectroscopy experiment (TOCSY) was first proposed by Braunschweiler and Ernst (15). This experiment has particular potential for protein NMR studies since it permits correlation of all protons within a scalar coupling network. Hence , a complete subspectrum can be obtained for every amino acid, making resonance assignments relatively easy (16). The pulse sequence for a conventional TOCSY experiment is shown in Figure 1.2B.

The most useful NMR experiment for the structural NMR spectroscopist involves without any doubt the through-space correlation. This correlation, named the Nuclear Overhauser effect (NOE) (17) , occurs between nuclei that are close in space. In simplest terms the magnitude of the NOE is proportional to $1/r^6$, where r is the distance between two nuclei. Due to this relation, the NOE fades quickly with distance, and NOE's are not normally observed between protons that are more than 5 Å apart. The pulse sequence for a conventional NOESY experiment is shown in Figure 1.2C. In the NOESY experiment, the mixing period occurs between the last two 90° pulses, during which cross-relaxation involving longitudinal magnetization occours.

Nuclear magnetic resonance (NMR) is the most useful method for peptide conformational analysis in solution. NMR yields information about chemical environments of individual nuclei, geometric relationship between nuclei, distances between nuclei and solvent accessibility of amide protons. When applied appropriately, NMR provides an exceptionally

powerful technique for determining the structure of the dominant conformers adopted by a peptide in solution. The determination of secondary structure of a peptide by NMR data can usually be accomplished via two methods, which may be combined to give an accurate determination of the secondary structural features of a peptide and/or protein. The first method involves interpretation of the NOE data, and is based on the fact that different secondary structures have different NOE patterns (different inter-proton distances). This approach has been discussed by Wüthrich (13), for example, the existence of α-helix is shown by the diagnostic α-β proton i, i+3 NOEs; while weak α-amide proton i, i+2 NOE usually indicate the existence of β-turn.

The second approach, known as the chemical shift index, relies solely on the chemical shifts of the backbone nuclei. The correlation between chemical shifts (H^{α} , H^{β}, $^{13}C^{\alpha}$, $^{13}C^{\beta}$,carbonyl ^{13}C) and the local secondary structure in proteins has been well studied and established (18). The α-proton net chemical shift with respect to the random structure was analysed in the structure determination. It has been recognized that in β sheets or extended regions α-proton resonance show downfield shifts , while in helices and turns, α-proton resonances show upfield shifts except those in position three of the type I β-turn.

The detection of intramolecular hydrogen bonds in peptides is important because when present, they can be expected to be one of the major noncovalent interactions providing conformational stability to a small peptide in solution (19). A useful NMR method for studying intramolecular hydrogen bonding is to follow the temperature dependence of the amide proton chemical shifts ($\Delta\delta/\Delta T$). In solvents such as water and ethanol, amide protons exposed to solvent can undergo fast exchange with the solvent protons, then the observed chemical shifts, δ_{obs} , are the weighed average of δ_{NH} (amide proton shift) and δ_{sol} (solvent proton shift) as given by the following equation:

$$\delta_{obs} = \alpha\delta_{NH} + (1-\alpha)\,\delta_{sol}$$

The terms α are 1-α are the fractions of the amide protons in the different states. Since δ_{sol} moves upfield as the temperature increases, so does the δ_{obs}. Normally, such solvent-exposed amide protons show values of $\Delta\delta/\Delta T$ from -6 to -10 ppb/°C. However, an amide proton involved in a stable intramolecular hydrogen bond, or one inaccessible to solvent for steric reasons, typically shows a reduced $\Delta\delta/\Delta T$ from 0 to -3 ppb/°C. Temperature coefficients in the intermediate range (from -3 to -6 ppb/°C) may indicate long noncoplanar intramolecular hydrogen bonds, or a conformational equilibrium between hydrogen-bonded and solvent exposed environments for the amide proton in question, or partial shielding of the proton from solvent.

In practice, the above procedures are combined and the schematic diagram of the steps used for the NMR structure determination in this thesis is shown in Figure 1.3.

Circular dichroism (CD) is the most extensively used technique for the exploration of peptide conformations in solution (20). The characteristic CD spectra of regular elements of secondary structure present in proteins and large polypeptides are well understood from experimental and theoretical work. The CD spectrum of a protein can be deconvoluted in terms of contributions from four components, namely the α-helix, sheet, reverse turn, and so called random coil. The diagnostic CD spectrum of α helix consists of a pair of strong negative ellipticities at 208 and 222 nm and a positive band near 190 nm (21). Theoretical analysis predict that The type I, II and III β-turn have similar CD spectra (22). The most common CD spectra for turns show a weak negative band between 220-230 nm, a positive band between 200 and 210 nm and a strong negative band between 180 and 190 nm. CD spectra of some β-turns may be quite different. For peptides that adopt relatively small

populations of helical conformations with dynamically frayed ends and, possibly, deviations from ideal helical structure, the presence of helix may be difficult to detect using CD(23,24).

Molecular Simulation

The computer simulation of protein/peptide structures should be viewed as a complement to, or an extension of, experimental work. However , the simulations should not be concerned exclusively with the structures but also with their functional characteristics, such as stability with respect to changes in pH, temperature, and solvent, the binding capability with respect to substrate, and catalytic properties (25).

In this thesis, the purpose of the simulation is aimed at modeling and refining the structure of the studied peptides based on the observed NMR data.

The structural properties for the flexible peptides can't be characterized by a unique structure, or a few discrete structures. The two main factors with a bearing on the approach to be adopted for the computer simulation are the temperature(T) and the time(t) dependence(26). In order to be completely safe, insofar as a comparison with experimental results, one should proceed, without doubt, with the consideration of both dependences.

The Monte Carlo (MC) search is a computing technique used to generate many random conformations, and the energy of these conformations consist of a Boltzmann distribution corresponding to the temperature being considered. In the MC simulations, the t-dependence is discarded. The generated structures can be optimized by energy minimization, and those with the lowest energy have the greatest chance to appear from the statistical point of view and are used as the model of the conformation of the studied peptide.

The MC simulation may yield a number of different conformations with comparable energy but they cannot reproduce the t-dependence implicit in the experiment. When the interpretation of experimental results is to be carried out at this level, molecular dynamics (MD) is the appropriate tool to be used (26). The dynamic behavior of the studied peptide can be evaluated by the trajectory of the atoms as a function of time, thus giving a description for what is happening to the peptide during the simulation.

The energy evaluation is carried out using an appropriate potential energy function (PEF). An energy minimization procedure is used to find the geometric conformation with lowest energy for the system under study. In the program BIOGRAF, used in this thesis, the DREIDING II force field (27) is applied. In this force field, the potential energy for an arbitrary geometry of a molecule is expressed as a superposition of the valence interaction that depends on the specific the structure and nonbonded interactions that depend only on the distance between the atoms. If constraints are used, the corresponding energy expression should be added and all the constraints use a harmonic force field and each is defined with an individual force constant (28). Distance constraints were used in this thesis, and the corresponding energy expression is:

$$V_{constraint} = (1/2)\, K_{constraint}\, (R-R_0)^2$$

The rigidity of each constraint can be controlled by changing the corresponding force constant. The value of the constraint force constant is usually about one order of magnitude larger than the normal value in order to keep the rigidity of the input constraints.

In the MD simulations the force, F_i, acting on the i-th particle, can be derived from the potential, V, (28). Because of the mass of atoms and the coupled nature (the dynamic equation of each particle depends on the position of all the particles in the system) of the Newtonian differential equations, molecular dynamics methods must solve the equations of

motion of the system with a fundamental time step of ~ femto-seconds (10^{-15} sec) (29). Thus, a typical simulation length is in the range of picoseconds (10^{-12} sec).

Reference

1. Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E. & Lerner, R. A., *Nature* 318, 480, 1985.
2. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E., *J. Mol. Biol.* 201, 161, 1988.
3. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E. & Lerner, R. A., *J. Mol. Biol.* 201, 201, 1988.
4. Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A. & Wright, P. E. , *J. Mol. Biol.* 226, 819, 1992.
5. Kim, P. S. & Baldwin, R. L., *Annu. Rev. Biochem.* 59, 631, 1990.
6. Jaenicke, R., *Biochemistry* 30, 3147, 1991.
7. Ptitsyn, O. B., *FEBS Letters* 285, 176, 1991.
8. Dyson, H. J., Lerner, R. A. & Wright, P. E., Annu. *Rev. Biophys. Biophys Chem.* 17, 305, 1988.
9. Richardson, J. S., *Adv. Protein Chem.* 34, 167, 1981.
10. Schulz, G. E. & Schirmer, R. H., *Principles of Protein Structure*, Springer-Verlag, New York, 1979.
11. Creighton, T. E., *Adv. Protein Chem.* 39, 125, 1984.
12. Dyson, H. J. & Wright, P. E., *Annu. Rev. Biophys. Biophys. Chem.* 20 519, 1991.
13. Wüthrich, K., *NMR of proteins and nucleic acids*. New York: John Wiley & Sons 1986.
14. Wemmer, D. E. & Reid, B. R., *Ann. Rev. Phys. Chem.* 36, 105, 1985.
15. Braunschweiler, L. & Ernst, R. R., *J. Magn. Reson.* 53, 521, 1983.
16. Bax, A., *Methods in Enzymology* 176, 151, 1989.
17. Noggle, J. H. & Schirmer, R. E., *The Nuclear Overhauser Effects*, Academic, New York, 1971.
18. Ösapay, K. and Case, D. A., *Journal of Biomolecular NMR*, 4, 215, 1994.

19. Hruby, V. J., "*Chemistry and Biochemistry of Amino Acids. Peptides and Proteins*" Weinstein, B. Ed., Vol. 3, Marcel Dekker, New York, 1974.

20. Bayley, P., *An Introduction to Spectroscopy for Biochemists*, Brown, S. B. Ed., Academic Press. New York, 148, 1980.

21. Johnson, W. C., Jr. & Tinoco, I., Jr., *J. Am. Chem. Soc.* 94, 4389, 1972.

22. Woody, R. W., *Peptides, Polypeptides and Proteins*, Blout, E. R., Bovey, F. A., Goodman, M. & Lotan, N. Ed., Wiley, New York, 338, 1974.

23. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. & Wright, P. E., *J. Mol. Biol.* 226, 795, 1992.

24. Dyson, H. J., Sayre, J. R., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A. & Wright, P. E., *J. Mol. Biol.* 226, 819, 1992.

25. Fraga, S., Parker, J. M. R., *Amino Acids*, 7: 175,1994.

26. Fraga, S., Parker, J. M. R. and Pocock, J. M., *Lecture Notes in Chemistry.* 66, ISNB 0-387-60133-3, New York, 1995.

27. Mayo, S. L., Olafson, B. D. and Goddar III, W. A., *J. Phys. Chem.*, 94, 8897, 1994.

28. *BIOGRAF Reference Manual*, Version 3.2.1, 1993.

29. Kollman, P., *Ann. Rev. Phys. Chem.*, 38: 303, 1987.

Table 1.1 Dihedral angles of hydrogen bonded β and γ turns

Turn type \ Dihedral Angles	i+1		i+2	
	ϕ	ψ	ϕ	ψ
β-Turn				
Type I	-60	-30	-90	0
Type I′	60	30	90	0
Type II	-60	120	80	0
Type II′	60	-120	-80	0
Type III	-60	-30	-60	-30
Type III′	60	30	60	30
Type VIa (cis)	-60	120	-90	0
Type VIb(cis)	-120	120	-60	0
γ-Turn	70 to 85	-60 to -70		
Inverse γ-Turn	-70 to -85	60 to 70		

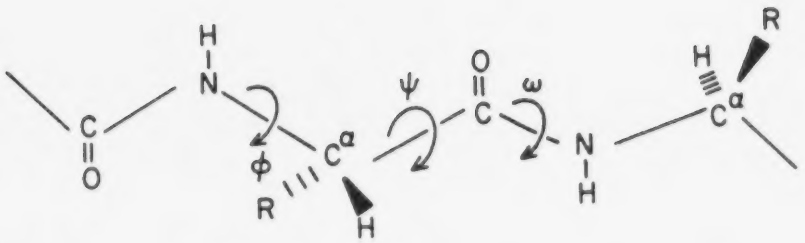


Figure 1.1 Standard nomenclature for the atoms and the torsion angles along a polypeptide chain

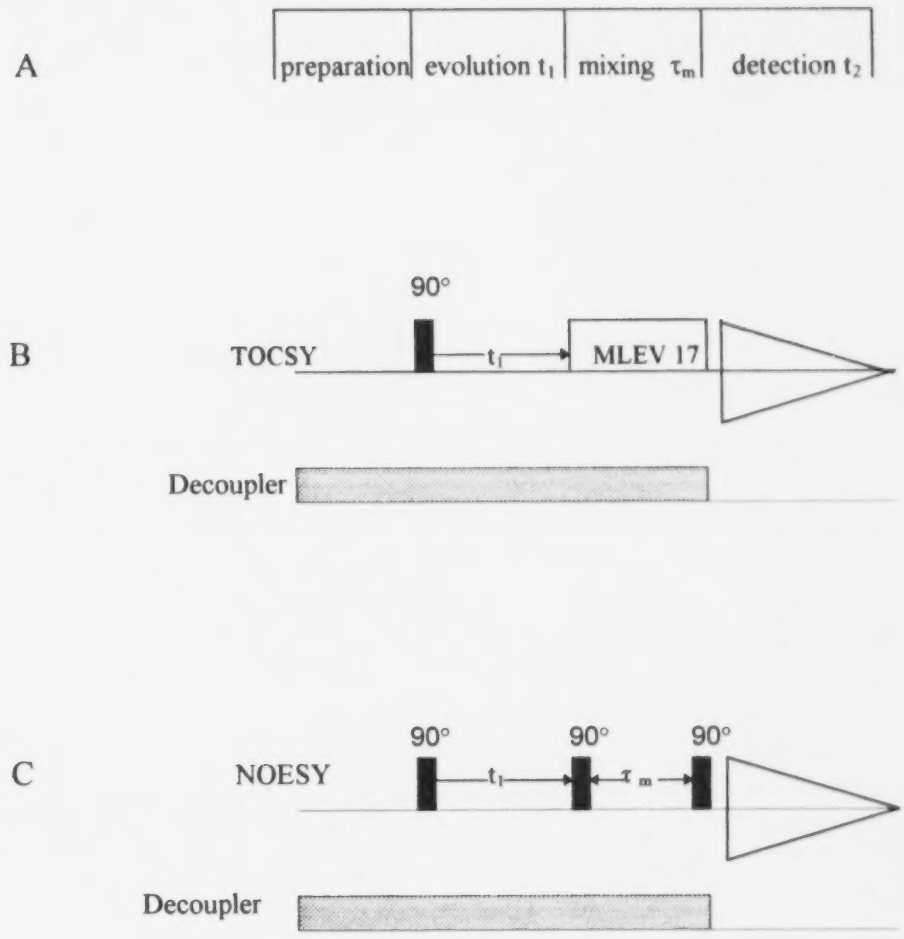


Figure 1.2 Pulse sequences for the two 2D NMR experiments used in this thesis

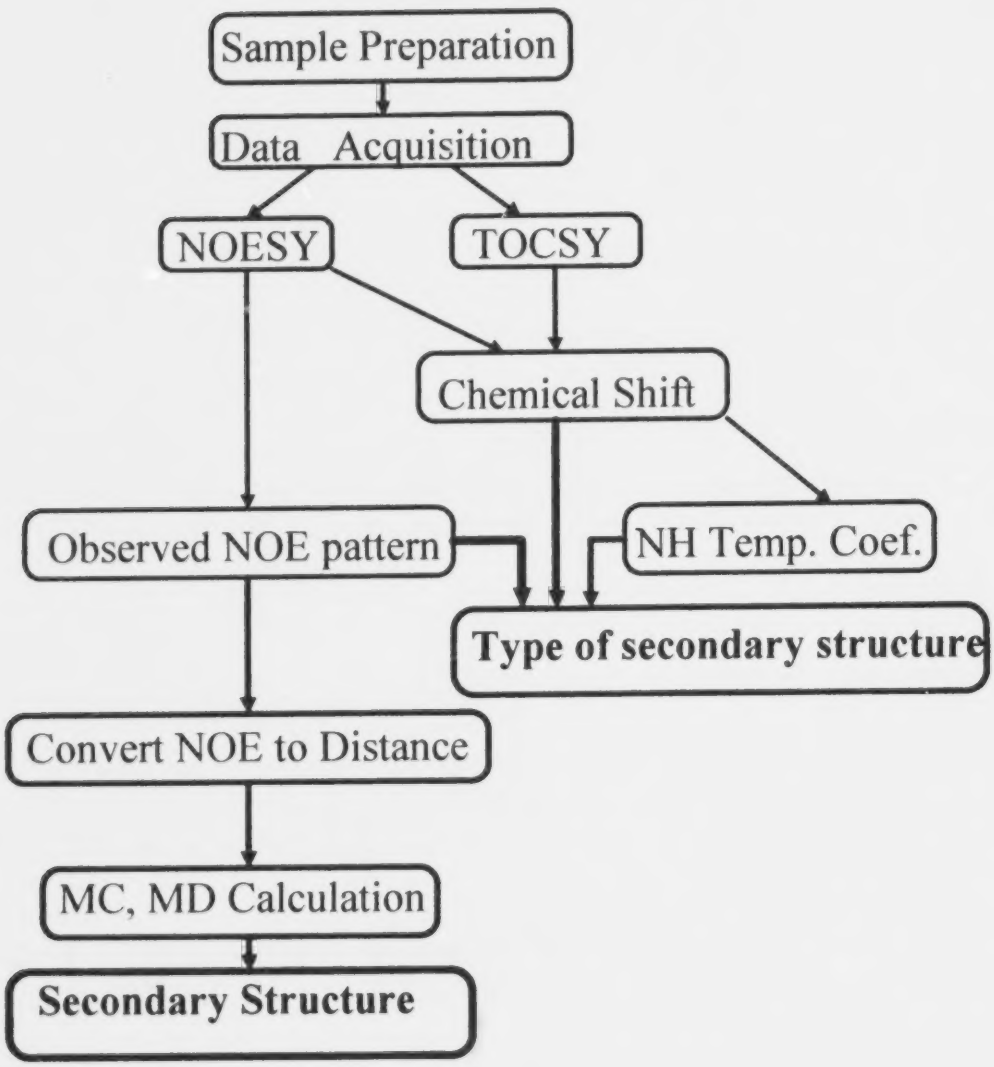


Figure 1.3 Schematic diagram of the steps used in this thesis for peptide structure determination by NMR

Chapter 2. Conformational Studies of Peptide Fragments from Decorin/ DS-PGII by ^{1}H NMR, CD and Molecular modeling

Background for this Project

Proteoglycans (PGs) are extracellular macromolecules which are present in and profoundly influence physical properties (e.g., resilience and distensability) of connective tissues. In non-cartilaginous soft connective tissues such as skin, tendon, ligament, etc., the predominant PGs are decorin and biglycan which consist of highly homologous protein cores of about 35 kDa (1) to which are attached glycosaminoglycan (GAG) chains of dermatan (DS) or chondroitin sulphate: one in decorin (2) and two in biglycan (3), close to the amino terminus. The GAGs are polyanionic and dominate the properties of the PGs carrying them. The small PGs may affect collagen fibril morphology and/or the organization of collagen fibrils into fibers and fiber bundles (4). Failure to efficiently synthesize GAG chains on the skin decorin protein core has been linked to the connective tissue derangement seen in progeroid syndrome (5).

Biosynthesis of GAG chains is initiated, probably in the rough endoplasmic reticulum, by the enzyme xylosyltransferase which transfers xylose from UDPto certain serines in the protein cores (6). The specificity of xylosyltransferase is incompletely defined but it is known that a xylosylated serine is always immediately N-terminal to glycine (most often) or to alanine (occasionally), some glycosylated sequences and predicted structure (7) in small DS-PGs are shown in Figure 2.1. Additional recognition factors must be important, however, because the sequences of both decorin and biglycan include several non-glycosylated Ser-Gly and Ser-Ala sequences(data is not shown here). So Similar uncertainty surrounds the control of O-glycosylation in glycoproteins (8). It

has been suggested, on the basis of predictions of secondary structure rather than on direct physical evidence, that serine should be in a β turn in order to be a candidate for glycosylation (9). The earliest sequence information on decorin was used to predict that residues 2 to 5 (Glu-Ala-Ser-Gly) would adopt a β-turn (10) and it was subsequently shown (2) that the single dermatan sulphate chain of skin decorin (then called "proteodermatan sulphate") was located on Ser^4. A preliminary circular dichroism study did not detect a β turn in a peptide comprising the amino-terminal 21 residues of bovine skin decorin (11) but in view of the evidence for proteolytic processing of the protein core subsequent to translation (12), this negative finding cannot be considered a valid test of the hypothesis. This is because the procore of decorin, the probable substrate for xylosyltransferase (11), carries an additional 14 amino acids between the suggested scissile bond for signal peptidase (13) and the N-terminus of the mature protein and could adopt a distinct secondary structure in this sequence.

In this thesis we report a study by NMR and circular dichroism spectroscopy of four peptides. The first peptide comprises the first 14 amino acids (Figure 2.2) of the mature decorin protein and the second has this same sequence extended in the N-terminal direction by the incorporation of an additional 10 amino acids derived from the procore of decorin (Figure 2.3). An additional 24-mer in which the methionine was oxidized to the sulfoxide, was also studied. The fourth peptide was the 14-mer with an acetyl group at the N-terminal end. This was included in an attempt to differentiate the effect of blocking the N-terminus from the effect of the additional N-terminal sequence in the 24-mer. The results indicate that an α-helix is present and could be a signal for xylosylation of serine in this small PG.

Materials and Methods

The 24 amino acid peptide CH3COQKGLFDFMLEDEASGIGPEEHFPENH$_2$ was purchased from the Alberta Peptide Institute. The 14-mer DEASGIGPEEHFPENH$_2$, the acetylated 14-mer CH$_3$CODEASGIGPEEHFPENH$_2$ and the oxidized 24-mer CH3COQKGLFDFM(O)LEDEASGIGPEEHFPENH$_2$ were provided by Dr. Paul G. Scott, they were synthesized on a Biolynx 4175 manual peptide synthesizer using FMOC amino acid derivatives (Pharmacia Canada Inc., Baie d'Urfe, Quebec). Peptides were purified by reverse-phase high performance liquid chromatography on a Vydac TP218 column (The Separations Group, Hesperia, CA). Purity was assessed by amino acid analysis and mass spectrometry. These four samples, 24-mer and oxidized 24-mer, 14-mer and acetylated 14-mer are abbreviated as NQ24, OQ24, D14 and AD14, respectively, in the following discussion.

Samples for CD experiments were prepared by dissolving purified peptides in water, and the pH was adjusted to 5.5 or 7.0 by adding 50 mM sodium phosphate buffer of that pH. The required volume of methanol was then added.

The CD spectra were recorded at room temperature on a JASCO JA20 spectropolarimeter with a thermostated cell holder. Mean residue ellipticities, $[\theta]_{mrw}$ (deg cm^2 dmol^{-1}), were calculated in the usual fashion using a mean residue weight of 108.1 for the 14-mer and 110.6 for the 24-mer, calculated from the composition of the peptides. Each spectrum is the average of 16 or 32 scans.

The samples used for ^{1}H NMR experiments were prepared by dissolving peptides in H_2O. The pH was then adjusted to 5.5 by adding diluted NaOH or HCl, followed by

CD_3OH to make the solvent a 60/40 methanol/water mixture. Before the NMR experiments, the sample was degassed with argon. After the 2D experiments, an aliquot of the OQ24 and the AD14 samples was diluted ten-fold with 60/40 CD_3OH/H_2O for the 1D experiments. These dilution experiments indicated that there was no aggregation.

All experiments were carried out on the Varian UNITY 500 NMR spectrometer using a proton-selective 5 mm probe with a 90° proton pulse length of about 9.0 μs at normal power levels. During the experiments the temperature was always controlled to ±0.1°C. For data collection and processing, VNMR software (versions 3.2 and 4.1) was used on the Sun Sparc 4/330 and Sparc 4/470 workstations, respectively. 2D spectra were acquired non-spinning, whereas for one-dimensional experiments the spinner was turned on. All chemical shifts are measured relative to the undeuterated fraction of the methyl group of CD_3OH at 3.30 ppm. All two-dimensional NMR experiments (14) were carried out in the phase-sensitive mode by using the hypercomplex technique (States-Haberkorn-Ruben method) (15). The decoupler and transmitter offsets were always set equal, a sweep width of 5000 Hz was used in both dimensions, and typically 2K (zero filling to 4K) data points in t_2 and 256 points (zero filling to 1K) experiments in t_1 were acquired. All two-dimensional experiments were preceded by 8 dummy scans but none were applied between individual t1-increments. The preacquisition delay alfa (16) was carefully and individually adjusted for every 2D experiment. First point distortions caused by analog filters (17) were empirically determined values. No further correction routines were necessary.

The TOCSY (18) experiments were carried out using the basic pulse sequence proposed by Bax (19). A 2 ms trim pulse preceded the 80 ms MLEV-17 spin-lock (field strength 7.0 kHz). The NOESY (20,21) experiments were carried out with the 90°-t_1-90°-t_m-90°-FID pulse sequence. For both experiments, water suppression was normally

achieved by decoupler saturation at all times except during the detection periods as shown in Figure 1.2. Approximately 90° shifted squared sine bell weighting functions were used in both dimensions for data processing.

Distances were calculated from NOESY spectra with a mixing time of 250 ms for all non-trivial cross peaks by integration of their 3D volumes in the following way: (a) cross peaks were integrated and carefully checked for contributions from baseline distortions; whenever possible, the volume was measured in both cross peaks, above and below the diagonal, and the two values were averaged to minimize slight phase imperfections; (b) where applicable, the volume was normalized for the contributions of more than two protons to the cross-peak by a factor $(2N_aN_b)/(N_a+N_b)$, where N_a and N_b are the number of protons responsible for the diagonal peaks a and b, respectively (22); (c) the $\beta\beta'$ interaction of the His and Phe protons was used as reference at 1.8 Å and (d) corrections for isomers arising from *cis*-proline were made.

The program BIOGRAF (version3.2.1, Molecular Simulations, Inc. Burlington, MA, USA) on the Sun SPARC 470 workstation (Sun OS 4.1.3 operating system) was used for all calculations. The program Hyperchem (Release 2, Autodesk, Inc.) was used for plotting the stereo structures. The DREIDING II force field was used for all calculations. For all energy minimizations, the conjugate-gradient method was used. 400-800 steps were typically needed for convergence (less than 0.1 kcal mol^{-1} RMS force). Distribution of charges was calculated by the Gasteiger method (23). All simulations were carried out for the molecule with all hydrogen atoms. Dielectric constant of vacuum (ε = 1) was used for all calculations. The molecular dynamic runs were performed with 1 fs steps at 300 K, unless another temperature is mentioned. The list of nonbonded interactions and hydrogen bonds was updated every 0.1 ps during the dynamic runs and every 50 steps during energy minimization. Cutoff distance for this list was 9 Å. In the

Monte Carlo searches, all dihedral angles were set to random values except angles in rings and ω dihedral angles in the backbone.

The acetylated 14-mer (AD14), and acetylated 24-mer with nonoxidized methionine residues (NQ24) were built and subjected to molecular modeling These constraints calculated from the NOESY spectra of OQ24 were used since these spectra were better resolved in comparison with that of NQ24, where heterogeneity caused by the *cis*-prolines produced more overlap.

Both structures were subjected to the Monte Carlo search: 50 conformers were generated and minimized without solvent shell. All atoms were set movable and all NOESY restraints were applied with a maximum force of 1000 k.cal/Å.mole. For both structures, the coordinates of the areas of interest (residues 11-14 for the 14-mer and 3-11 for the 24-mer) in the three lowest energy conformers were matched. Resulting structures were averaged and minimized with constraints. Minimized structures were solvated with water (diamond grid, 2.8 Å distance between solvent molecules, 2.6 Å inner cutoff of shell from peptide, 6.6 Å outer cutoff) and the whole system was minimized again. The unsolvated 24-mer and solvated 14-mer and 24-mer were subjected to 100 ps dynamic runs. The first 80 ps were constrained, the last 20 ps were unconstrained. Every 1ps the structure was written to the trajectory. Coordinates of 20 conformers from the last 20 ps were averaged (peptide only), resolvated where dynamic runs were performed with the solvent shell, and minimized. Backbone dihedral angles of the resulting structures are then evaluated.

The structure of the nascent helix was also tested by the following procedure: The N-terminal part of the 14-mer (Residues 1 to 10) was folded into a α-helix with backbone torsion angles ϕ=-57°, ψ=-47° and ω=-180° and the resulting structure was subjected to

constrained minimization. Then a constrained dynamic run of overall length 20 ps was performed. Every 1 ps the structure was minimized without constraints and written to the trajectory file.

Results

The Circular Dichroism spectra in Figure 2.4 show that both the unacetylated 14-mer and the acetylated 24-mer possessed little if any regular secondary structure in a purely aqueous solvent. Including methanol at 60 or 70% (v/v), however, induced some secondary structure, more marked in the 24-mer than the 14-mer. Replacing the methanol with an equal volume of trifluoroethanol gave identical spectra, as did altering the pH from 5.5 to 7 and increasing the concentration from 80uM to 800 uM (data not shown) for 24-mer.

The resonance assignments listed in Tables 2.1-2.4 of the four studied peptides are based on TOCSY experiments and sequential NOE's of the type αH(i)-NH(i+1), NH(i)-NH(i+1) and βH(i)-NH(i+1) (24).

Additional assignment problems arose from the presence of the *cis*-isomers of the two prolines. The *cis* form allows much closer contacts between αH_i and αH(i+1) and between NH(i) and αH(i+1), whereas the *trans* form favors short distances between αH(i) and δH_2 (i+1) and between NH(i) and δH_2 (i+1). The correlation peaks between 2H and 4H and between βH and 4H of the histidine residue also help to identify the existence of the isomers. The presence of *cis* isomers for the two prolines was unequivocally shown by NOESY spectra (see Figures 2.5). For the two 14-mers studied, the strong cross peaks between αH of Phe^{12} and δH_2 of Pro^{13}, between αH of Gly^{7} and δH_2 of Pro^{8}, between δH_2 of Pro^{13} and NH of Glu^{14} and between δH_2 of Pro^{8} and NH of

Glu[9], show that these two prolines are mainly in the *trans* conformation. The two sets of cross peaks between αH of Phe[12] and αH of Pro[13] and between αH of Pro[13] and NH of Glu[14] show the presence of the *cis* form of the Pro[13]. The small proportion of *cis* isomer for Pro[8] was demonstrated by the weak interactions between αH_2 of Gly[7] and αH of Pro[8]. Weak intra residue interactions between αH_2 and NH of Gly [7] also could be observed; however, the low population precluded obtaining further NOE interaction patterns for this *cis*-isomer. No further interactions of the α-protons between Gly[7] and Pro[8] and between Phe[12] and Pro[13] were detected, so three isomers arising from *cis* proline were found. They were: *trans*(Pro[8])-*trans*(Pro[13]); *trans*(Pro[8])-*cis*(Pro[13]) and *cis*(Pro[8])-*trans* (Pro[13]) (abbreviated as TT, TC and CT respectively).

Resonance heterogeneity arising from *cis* Pro [8] and *cis* Pro [13] can be found in the region from Ile[6] to the C-terminal NH and in the region from Gly [9] to the C-terminal NH respectively. Some residues between these two prolines suffer the effects from both *cis* prolines, and their resonances are tripled. Some spectra with doubled or tripled resonances are shown in Figures 2.6 and 2.7.

The stereo structure of the *cis* and *trans* proline and the assignment of ring protons is shown in Figure 2.8. However, for the TC isomers, not only are the assignments unambigous due to the well-resolved NOESY and TOCSY spectra , but the spectra also allow the stereo specific assignments of the ring protons of proline (25) to be evaluated (Figure 2.9 and Tables 2.1-2.4). Due to its low population, the assignments of the CT isomer are incomplete

This discussion also holds true for the two 24-mers studied, adding 10 amino acids at the N-terminal has no effects on the isomerization of the two prolines.

By comparing the integrals of the terminal NH in the one-dimensional spectra, the TT, TC and CT isomers were found to be present in approximate 68:25:7 proportions for all the four peptides studied. An observed medium-range NOE between αH of Phe^{12} and NH of Gly^{14} (see Figure 2.10) suggests the presence of a turn. Molecular modeling shows a type VIb β turn, where the *cis* Pro^{13} would be in the (i+2) position in the turn. The three isomers discussed above as well as the type VIb β turn in the TC isomer exist in the same proportion in all four peptides studied. No further medium or long-distance NOE's were detected for the CT isomer due to its low abundance.

The same NOE patterns were observed for the peptides AD14 and D14: continuous dNN(i,i+1) connectivity, consecutive weak dαN(i,i+2) connectivity (see Figure 2.10) and the lack of (i)-(i+3) medium range NOE's, show that these two peptides have a similar conformation, that is a nascent helix (26), which is characterized by an ensemble of turn-like structures in the early stages of helix formation (27). Similar sequences of i, i+2 NOE connectivity were observed by Dyson et al. for some short peptides(26,27). This structural element was named the "nascent helix" and it was assumed to consist of rapidly interconverting turns. Our nascent helix in the 14-mer is a typical case of rapidly interconverting turns. No i, i+3 connectivities characteristic of an α-helix were found and the absence of the helical fragments is also seen in the CD spectra.

There are close connections between chemical shifts and the local secondary structure (28,29). The net structural shifts obtained by comparing the chemical shift with the random structure (24) for the peptides AD14 and D14 are shown schematically in Figure 2.11. The closely similar net shift patterns indicate similar structures for D14 and AD14. The more negative value of D1 of peptide D14 is attributed to the effect of the positive charge at the N-terminus. The similarity of the NH temperature coefficients (see

Tables 2.1-2.4) also confirms the same structure for each pair of peptides. This agrees with the CD results which were identical for D14 and AD14 (not shown).

The presence of the sulfoxide in the Met residue in peptide OQ24 splits the peaks of Met^{8} into two signals of equal intensity, indicating that equal amounts of the R and S sulfoxide confurations were formed(30). The NH peaks of Leu^{9} and Gly^{10} were also doubled. However, the NMR studies show that the sulfoxide group has no effect on the OQ24 peptide conformation. The peptides NQ24 and OQ24 have the same conformation. The medium range (i)-(i+3) NOE's, strong NH-NH interactions (see Figures 2.12-2.14), low temperature coefficients(see Tables 2.3-2.4) and the α-proton shifts (see Figure 2.15), indicate that the peptides OQ24 and NQ24 adopt an α-helical structure from Lys^{2} to around Gly^{15}. The low temperature coefficients from Asp^{6} to Gly^{15} indicate that the NH protons are protected from solvent and probably hydrogen-bonded. It has now been well established that the αH protons of residues in an α-helical conformation experience an upfield shift with respect to random coil chemical shift values. For peptides that exhibit exclusively random coil and α-helical conformations, with a rapid interconversion between both states, one can expect that the αH proton of each residue will suffer an upfield shift proportional to the amount of time the residue spends in an α-helical conformation (31).

A semiquantitative estimation of the helical content of peptides can be obtained by dividing the average α proton shift by 0.35, the average upfield α proton shift observed in the amino acid residues of proteins in an α-helical conformation (32). Using this method, the 24-mer peptides are estimated to adopt about 65 % helical conformation which starts from Lys^{2} and ends around Gly^{15}. Considering the existence of the isomers of the peptides studied and the nature of the CD spectra as was discussed in Chapter 1, CD is not a suitable approach to estimate the helical content for the system we studied.

The observed NOE's for the AD14TT, AD14TC and OQ24TT isomers are summarized schematically in Figure 2.16; and the calculated distance constraints which were used in the modeling are listed in Table 2.5 and Table 2.6 for the AD14 TC and OQ24 TT isomers, respectively.

All of the above data have been correlated with molecular modeling. As was mentioned above, both measured 14-mers give a similar set of NOESY cross peaks useful for the calculation of distance constraints; in the case of the 24-mers, the situation is similar. Furthermore, the C-terminal part of all peptides shows the same NMR patterns (NOESY, chemical shifts, temperature dependencies, *cis-trans* isomerization); therefore, the proposal that the C-terminus of all peptides has the same conformation is reasonable. Two structures were built for evaluation by molecular modeling: 1) AD14 with a *cis* bond between Phe^{12} and Pro^{13} for the evaluation of the nascent helix in the N-terminal part of the peptide and the nature of the turn at C-terminus. The nascent helix is present in both 14-mers, and the C-terminal turn is present in all peptides with a *cis* proline in position 13 for the 14-mers or position 23 for the 24-mers; 2) The acetylated 24-mer(NQ24) with all peptide bonds *trans* for the determination of the structure of the α-helix at the N-terminus of both 24-mers.

Our modeling study for the nascent helical structure of the 14-mers started with constrained minimization of randomly generated conformers. Some helical structures were obtained, but these structures quickly unfold when constraints are removed. Unfolding leads to structures with most i, i+2 distances observed in the NOESY spectra between 4 and 7 Å. Also, a dynamic run starting from the α-helix (see Experimental) shows an instability of the whole system. From 20 conformers, 11 relax after removing the constraints to the fragments of an α-helix, and 9 relax to structures containing 1 or 2 ideal or non ideal turns, mainly type II β-turns. These calculations are in agreement with

Dyson et al (26, 27, 33), namely, that no stable structure is characterized by consecutive i, i+2 NOE connectivities, and these connectivities arise from random structures containing some population of turn in various positions. Table 2.7, column A, gives the backbone dihedral angles for the averaged structure from the last 20 ps of the 100 ps dynamic run(see experimental). It is seen that this structure relaxes to a random coil with a nonideal type I β-turn in positions Ser^4 to Gly^7.

The same modeling of the AD14 reveals the nature of the C-terminal turn in the conformations containing the *cis* peptide bond between Phe^{12} and Pro^{13}. Because of the low population of *cis* peptide, no additional constraints to the basic i, i+2 were available, and the Monte Carlo search with constrained minimization led to various turns. The average of the dynamic run starting from the three best conformers led to the nonideal typeVIb β-turn (12) (see Table 2.7 , column A) as the most probable or most populated conformation for the C-terminus for our peptides containing *cis*-proline. Figure 2.17 represents the match of the last 20 conformers from the 100 ps dynamic run for the peptide AD14 with *cis* proline in position 13.

The average conformer from the Monte Carlo search of the 24-mer was subjected to dynamic runs without solvent as well as in the solvent (water) shell. Figure 2.18 shows the averaged minimized structure for the 24-mer with all bonds *trans* modeled without solvent; the backbone data for this structure are listed in Table 2.7, column B. Figure 2.19 represents the stereo view of the average minimized structure of the peptide after modeling with the shell of solvent; Backbone data for this structure are in Table 2.7, column C. Both structures exhibit a helical structure from Lys^2 to Gly^{15}. Analysis of the chemical shifts of the α protons and NOE connectivity shows a continuous helical structure; the temperature dependencies of the NH protons suggest possible disruptions of helix at residues 8 and 11. The structure modeled without solvent shell exhibited

disruption of the helical structure at residue 13. This is probably caused by an over-estimation of repulsion between charges in the preceding residues during the calculation "in vacuo". When the dynamic run is performed with peptide in the water shell, the structure of the helix is improved (see data in Table 2.7, compare columns B and C). Finally, dihedral angles at the C terminal end of the helix (involving residues 14-17) are similar to those for a type I β-turn.

Discussion

A comparison of the data for the acetylated 14-mer(AD14) with the unacetylated 14-mer(D14) indicates a nascent helix in both peptides. This is consistent with the N-terminal capping preferences for amino acids(34) in α-helical peptides. For example, for the acetyl group the helix fraction is 0.578, whereas when the charged Asp is the N-terminal amino acid with a charged NH_3^+ N-terminus, the fraction of helix for the same control peptide is 0.527(34)

At the C-terminal end of the peptides, the temperature coefficients and α-proton shifts show that the sequence Ile-Gly-Pro-Glu-Glu-His-Phe-Pro-Glu is of the same conformation in all four peptides studied. The overlap of the spectra and the resonance heterogeneity caused by the prolines prevent further conclusions. The unique ability of the proline to form *cis* peptide bonds and undergo *cis-trans* isomerization (35,36) is well known. The significance of the study of the isomerization of the proline is two-fold: first the isomerization of proline induces resonance heterogeneity for amino acids adjacent to it, on the other hand, the *cis* and *trans* isomerization forms differ in their solution conformation and are of biological significance(36). In this work, the peptides have two proline residues, and statistically, four possible isomers could be produced, while, in fact only three were found.

The *cis-trans* ratio of the *X*-*L*-Pro peptide bond is strongly influenced by *X* , the residue preceding the proline. It was found that the relative abundance of *cis* proline increases when one goes from *X*= Gly to *L*-Ala to *L*-Leu to *L*-Phe, and this different behavior was interpreted in terms of a single structural parameter, i.e., the bulkiness of the hydrophobic side chains (37). An aromatic residue preceding proline predisposes it to the formation of a large proportion of *cis* isomer in solution (38). In the four peptides we studied, the two residues preceding the two prolines are Gly and Phe, and the abundance of the *cis* proline is 7% and 25% respectively. Our results agree with the above discussion. The *cis-cis* (CC) isomer was not detected and we propose that the behavior (*cis-trans* conversion) of each proline is independent of each other, so the chance to form the CC conformation is quite small (7% * 25% = 2%).

The addition of the N-terminal 10 amino acids in the 24-mers forms a complete helix ranging from the N-terminal end Lys^{2} to Gly^{15}. Consequently, the region of interest Asp^{11}-Glu^{12}-Ala^{13}-Ser^{14}-Gly^{15} is part of the α-helix, and the N-terminal 10 amino acids promote the α-helix character for this region of the peptide that contains the Ser that is xylosylated in decorin. Bourdon et al(39) carried out a kinetic analysis of the xylose-acceptor efficiencies (Vmax/Km) of two peptides: FMLEDEASGIGP and CDEASGIGPEVPDDRD related to human decorin and several variants of these. Both peptides were acceptors, indicating that they contain the minimum recognition sequence for xylosyltransferase activity, which may be SGXG(39). The N-terminal cysteine in CDEASGIGPEVPDDRD (which is not present in decorin) may be analogous to the N-acetyl group in the 14-mer studied here. It would be of interest to study the solution conformations of these peptides to determine whether they adopt the (nascent) helical structures described here.

While the adoption of an α-helical conformation by the N-terminal segment of the 24-mer could have been predicted from the amino acid sequence by the method of Chou and Fasman(40), the inclusion of the first 5 amino acids(Asp-Glu-Ala-Ser-Glu) of the mature proteoglycan in this helix is somewhat unexpected because this segment was originally predicted to adopt a type I β-turn (10), and these are usually considered to disrupt helices (40). In either a β-turn or an α-helix, the serine (Ser^{14}), which in decorin becomes xylosylated, would be exposed on the outside of a turn. This may be one structural factor controlling glycosylation enzyme activity.

The N-terminal portion of the helix in the 24-mer shows a peak of hydrophobicity and amphipathicity, detected by the sequence helical hydrophobic moment method of Eisenberg(41, data not shown), centered around Phe^{5}. Amphipathic helices are found on the surfaces of globular proteins and are common in α/β proteins, the structural class to which decorin and the other small proteoglycans almost certainly belong(10). This helix could thus serve to position the consensus xylosylation sequence for accessibility to xylosyltransferase.

Conclusions

In methanol/water mixtures, Asp^{1}-Glu^{2}-Ala^{3}-Ser^{4}-Gly^{5} (DEASG) at the N-terminal end of the 14-mer peptide forms a nascent helical structure. Acetylation of the N-terminus has no effect on the structure. The addition of a further 10 amino acids to the N-terminus forms an helical structure which begins at the N-terminus Lys^{2} and ends at or near Gly^{15}. The results indicate that an α-helix is present and suggest that this could be a signal for xylosylation of serine in this small PG.

Three isomers induced by the *cis-trans* isomeration of the two prolines (Pro^{8}, Pro^{13} in the 14-mers, and Pro^{18}, Pro^{23} in the 24-mers) in the four studied peptides were detected from TOCSY and NOESY ^{1}H NMR spectra. The *trans-trans, trans-cis* and *cis-trans* isomers exist at a ratio of about 68:25:7 in the methanol/water mixture. The resonance heterogeneity induced can be observed as far as 4 amino acid residues away from proline. A type VIb β-turn has been observed involving the C-terminal *cis* proline in the sequence His -Phe -Pro -Glu based on the observed medium range NOE between αH of Phe^{12} and NH of Glu^{14} in the two 14-mers and one between αH of Phe^{22} and NH of Glu^{24} in the two 24-mers and the modeling results of the peptide AD14.

Futher work proposal

1. In this thesis, we have investigated the conformation for one of many peptide fragments involving the xylosyltransferase process. Based on the experimental results, the helical structure is proposed to be a recognition factor for the peptide studied. However, this can not exclude other possible factors, and the α helix, turn, both helix and turn, or others may be a recognition factor. Obviously, to finally determine the recognition factor for the xylosyltransferase process from the point of the view of structure, further investigations on other peptides involving this process should be carried out. This project is not the end but just the beginning. Considering the helical structure detected in this paper, in future work, the peptide fragments should be chosen long enough to maintain the stability of the helical structure.

2. This work has focused on the conformational study of the peptide fragments containing the amino acids Ser-Gly which are involved in the xylosyltransferase reaction, while the conformational study of the peptide fragments containing the Ser-Gly or Ser-Ala fragment which is not involved in the xylosyltransferase reaction is of course necessary.

3. Three isomers induced by *cis* proline in the four studied peptides were found. More work can be carried out on this interesting part, namely: (a) Stability of such isomers can be studied in different solvents, at different pH and temperature. (b) A ^{13}C NMR study would help to obtain more detailed structural information for these three isomers and will also allows us to collect some chemical shift data.

Reference

1. Fisher, L. W., Termine, J. D., and Young, M. F., *J. Biol. Chem.* 264, 4571, 1989.
2. Chopra, R. K., Pearson, C. H., Pringle, G. A., Fackre, D. S., and Scott, P. G., *Biochem. J.* 232, 277, 1985.
3. Neame, P. J., Choi, H. U., and Rosenberg, L. C., *J. Biol. Chem.* 264, 8653, 1989.
4. Scott, J. E., Orford, C. R., and Hughes, E. W., *Biochem. J.* 195, 573, 1981.
5. Quentin, E. A., Gladen, L. R., and Kresse, H., *Proc. Natl. Acad. Sci.* USA 87, 1324, 1990.
6. Hoffman, H. P., Schwartz, N. B., Rodén, L. and Prockop, D. J., *Connect. Tiss. Res.* 12, 151, 1984.
7. Wilmot, C. M. and Thornton, *J. M. J. Mol. Biol.*, 203, 221-232, 1988.
8. Wilson, I. B., Gavel, Y. and Heijne, G., *Biochem. J.* 275, 529, 1991.
9. Aubert, J. P., Biserte, G., and Loucheux-Lefevre, M. H., *Arch. Biochem. Biophys.* 175, 410, 1976.
10. Pearson, C. H., Winterbottom, N., Scott, P. G., and Fackre, D. S., *Biochem. Soc. Trans.* 11, 747, 1983.
11. Scott, P. G., In Dermatan Sulphate Proteoglycans: *Chemistry, Biology, Chemical Pathology.* Edited by Scott, J. E., Portland Press, London, p. 81, 1981.
12. Sawhney, R, S., Hering, T. M. and Sandell, L. J., *J. Biol. Chem.* 266, 9231, 1991.
13. Krusius, T and Ruoslahti, E., *Proc. Natl. Acad. Sci.* USA 83, 7683, 1986.
14. Ernst, R. R., Bodenhausen, G. and Wokaun, A. *In Principles of Nuclear Magnetic Resonance in One and Two Dimensions.* Clarendon Press, Oxford, 1987.
15. States, D. J., Haberkorn, R. A., and Ruben, D. J., *J. Magn. Reson.* 53, 528, 1983.
16. Varian NMR Spectrometer Systems. *System Operation Manual.* VNMR. Pub. No. 87-195100-00. Rev. D0192.
17. Otting, H. G., Widmer, G. W. and Wüthrich, K. *J. Magn. Reson.* 66, 187, 1986.
18. Braunschweiler, L. and Ernst, R. R., *J. Magn. Reson.* 53, 521, 1983.

19. Bax, A. and Davis, D. G., *J. Magn. Reson.* 65, 355, 1985.

20. Jeener, J., Meier, R., Bachmann, P. and Ernst, R. R. *J. Chem. Phys.* 71, 4546, 1979.

21. Kumar, A., Ernst, R. R. and Wüthrich, K., *Biochem. Biophys. Res.* Commun. 95, 1, 1980.

22. Williamson, M. P. and Neuhaus, D., *J. Magn. Reson.* 72, 369, 1987.

23. Gasteiger, J. and Marsili, M., *Tetradhedron* 36, 3219, 1980.

24. Wüthrich, K., *NMR of Proteins and Nucleic Acids.* Wiley-Interscience, New York, 1986.

25. Pogliani, L. and Ellenberger, M. *Organic Magnetic Resonance*, Vol. 7, 1975.

26. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E. and Lerner, R. A. *J. Mol. Biol.* 201, 1988.

27. Dyson, H. J. and Wright, P. E., *Annu. Rev. Biophys. Biophys. Chem.* 20, 519, 1991.

28. Wishart, D. S., Sykes, B. D. and Richards, F. M., *J. Mol. Biol.* 222, 311, 1991.

29. Ösapay, K. and Case, D. A., *J. Biomol. NMR* 4, 215, 1994.

30. Markley, J. L. and Kainosho, M., *In NMR of Macromolecules. A practical Approach.* Edited by Roberts, G. C. K., Oxford University Press. 1993.

31. Josep R., Francisco J. B., Bostjan K., Martha D. B. and Lila M. G., *Biochemistry*, 32 4881, 1993.

32. Annet, R., Garry, W. B., and Robert, J. C., *Biochemistry*, 34, 7401, 1995.

33. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. and Wright, P. E., *J. Mol. Biol.* 226, 795, 1992.

34. Doig, A. J. and Baldwin, R. L., *Protein Sci.* 4, 1325, 1995.

35. Amodeo, P., Antonietta, M., Morelli, C. and Motta, A., *Biochem.* 33, 10754, 1994.

36. McInnes, C., Kay, C. M., Hodges, R. S. and Sykes, B. D., *Biopol.* 34, 1221, 1994.

37. Grathwohl, C. and Wüthrich, K., *Biopolymers* Vol. 15, 2025-2041, 1976.

38. Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H-C., Lerner, R. A. and Wright, P. E. J. Mol. Biol. 226, 819, 1992.

39. Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B. and Ruoslahti, E., *Proc. Natl. Acad. Sci.* USA 84, 3194, 1987.

40. Chou, P. Y., *In Prediction of Protein Structure and Principles of Protein Conformation.* Edited by Fasman, G. D., Plenum Press, New York, p. 549, 1989.

41. Eisenbert, D., Weiss, R. M. and Terwilliger, T. C., *Nature* 299, 371, 1982.

Table 2.1 ^{1}H NMR Resonance Assignments [a] and NH Temperature Coefficients [b] of the peptide D14[c].

Residue	NH	Hα [d]	Hβ [d]	Hγ [d]	Other H [d]	NH Temp. Coef. (-ppb/°C)
D1		4.18	2.91;2.86			
E2	9.23	4.20	2.07;1.97	2.32;2.32		3.9
A3	8.35	4.28	1.42			2.8
S4	8.24	4.37	3.93;3.88			3.7
G5	8.25	4.04;3.86				2.2
I6	7.96	4.25	1.82	1.43;1.10	γCH$_3$:0.89; δCH$_3$:0.82	3.2
G7	8.51	4.15;4.05				7.1
P8		4.37	2.25;2.25	2.00;2.00	δδ′:3.71;3.60	
E9	8.81	4.16	1.92;1.92	2.25;2.25		3.7
E10	8.16	4.16	1.87;1.87	2.15;2.15		2.9
H11	8.27	4.61	3.10;3.01		2H:8.45; 4H:7.15	5.4
F12	8.47	4.81	3.18;2.90		7.29 (ave. ring)	6.8
P13		4.37	2.25;2.25	1.98;1.98	δδ′:3.72;3.55	
E14	8.58	4.22	2.07;1.95	2.31;2.31		5.8
NH_2	7.63;7.14					4.8;5.9
I6 CT	7.93	4.25	----;----	----;----		---
G7 CT	----	4.03;3.69				5.5
P8 CT		4.56	----;----	----;----		
E9 CT	8.76	4.25	1.90;1.90	2.24;2.24		4.3
H11CT	8.47		3.08;3.00		----;----	5.3
F12CT	8.55		3.17;2.88		----	6.9
NH_2 CT	7.64;7.14					
E9 TC	8.83	4.22	1.92;1.92	2.26;2.26		3.9
E10 TC	8.19	4.23	1.92;1.92	2.20;2.20		3.2
H11TC	8.38	4.64	3.19;3.08		2H:8.51; 4H:7.26	5.1
F12 TC	8.38	4.63	3.05;2.92		7.24 (ave. ring)	7.0
P13 TC		3.78	1.86[1];1.35[2]	1.66[1];1.63[2]	δδ′:3.52[1]; 3.29[2]	
E14 TC	8.83	4.12	2.02;2.02	2.31;2.31		5.8
NH_2 TC	7.83;7.13					5.8;5.9

[a] Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30ppm downfield from TMS.

[b] NH temperature coefficients are the result of a linear regression analysis of the chemical shifts at 5, 10, 15, 20 and 25 °C TOCSY spectra.

[c] The data were recorded at 500 MHz in CD_3OH/H_2O (60/40 v/v) at a concentration of 2.5mM. The pH was 5.5 in water, the temperature was 5 ±0.1 °C.

[d] Pairs of geminal protons x (= β, γ, δ, ε and α(G)) are not assigned stereospecifically, except those for *cis* Pro[13]. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x′

Table 2.2 ^{1}H NMR Resonance Assignments [a] and NH Temperature Coefficients [b] of the peptide AD14 [c].

Residue	NH	Hα [d]	Hβ [d]	Hγ [d]	Other H [d]	NH Temp. Coef. (-ppb/°C)
CH_3CO					CH_3CO: 2.01	
D1	8.31	4.59	2.69;2.62			5.6
E2	8.86	4.18	2.08;1.96	2.31;2.31		5.0
A3	8.41	4.30	1.41			4.3
S4	8.08	4.34	3.89;3.89			2.5
G5	8.26	4.00;3.86				2.7
I6	7.88	4.22	1.85	1.43;1.11	γCH_3:0.88;δCH3:0.82	3.3
G7	8.45	4.18;4.02				6.2
P8		4.37	2.25;2.25	1.97;1.97	δδ':3.70;3.61	
E9	8.74	4.17	1.92;1.92	2.24;2.24		3.7
E10	8.14	4.16	1.86;1.86	2.15;2.15		2.9
H11	8.26	4.60	3.10;3.00		2H:8.48;4H:7.18	5.6
F12	8.46	4.82	3.17;2.90		7.28 (ave. ring)	7.3
P13		4.38	2.25;2.25	1.97;1.97	δδ':3.72;3.55	
E14	8.55	4.23	2.07;1.95	2.31;2.31		6.3
HH_2	7.62;7.13					4.7;6.0
G7 CT	8.27	4.04;3.68				6.0
P8 CT		4.58	----;----	----;----	----;----	
E9 CT	8.72	4.25	1.92;1.92	2.24;2.24		4.2
H11CT	8.47	4.62	3.11;3.00			6.5
F12 CT	8.54	4.82	3.17;2.89			7.8
NH_2 CT	7.66;7.13					4.5;5.7
E9TC	8.77	4.23				4.2
E10 TC	8.17	4.23	1.88;1.88	2.21;2.21		3.0
H11 TC	8.36	4.65	3.21;3.06		2H:8.54;4H:7.27	5.0
F12 TC	8.36	4.63	3.06;2.92		7.23(ave. ring)	7.3
P13 TC		3.80	1.87 ^{1};1.38 ^{2}	1.67 ^{1};1.64 ^{2}	δ:3.52^{1};3.29^{2}	
E14 TC	8.79	4.13	1.97;1.97	2.27;2.27		6.2
NH_2 TC	7.80;7.10					5.9;5.7

[a] Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30ppm downfield from TMS.

[b] NH temperature coefficients are the result of a linear regression analysis of the chemical shifts at 5, 10, 15, 20 and 25 °C TOCSY spectra.

[c] The data were recorded at 500 MHz in CD_3OH/H_2O (60/40 v/v) at a concentration of 2.9 mM. The pH was 5.5 in water, the temperature was 10 ±0.1 °C.

[d] Pairs of geminal protons x (= β, γ, δ, ε and α(G)) are not assigned stereospecifically, except those for *cis* Pro[13]. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x'

Table 2.3 ^{1}H NMR Resonance Assignments [a] and NH Temperature Coefficients [b] of the peptide OQ24 [c].

Residue	NH	Hα [d]	Hβ [d]	Hγ [d]	Other H [d]	NH Temp. Coef. (-ppb/°C)
CH_3CO					CH_3CO : 2.01	
Q1	8.35	4.27	2.03;1.93	2.31;2.31	δNH_2 : 7.70;6.68	5.2
K2	8.50	4.30	1.72	1.44;1.38	δδ′:1.59;1.59	5.0
					εε′: 2.86;2.84	
G3	8.78	4.12;3.87				7.6
L4	8.29	4.10	1.50;1.45	1.35	δCH_3:0.86;0.79	7.7
F5	8.29	4.48	3.15;2.95		7.24 (ave. ring)	9.3
D6	7.86	4.44	2.60;2.60			+0.6
F7	8.04	4.42	3.17;3.05		7.25 (ave. ring)	3.0
M8I [e]	8.32	4.36	2.25 [f]	2.89 [f]	εCH_3 :2.68	3.2
M8II [e]	8.30	4.33	2.29 [f]	2.96;2.94	εCH_3 :2.79	2.9
L9I [e]	8.18	4.18	1.69;1.69	1.54	δCH_3 :0.88;0.85	2.8
L9II [e]	8.16	4.18	1.69;1.69	1.54	δCH_3 :0.88;0.85	2.7
E10I [e]	8.37	4.12	2.05;2.05	2.35;2.35		0.3
E10II [e]	8.32	4.11	2.052.05	2.35;2.35		0.1
D11	8.32	4.49	2.73;2.73			5.4
E12	8.42	4.14	2.09;2.09	2.44;2.33		1.6
A13	8.27	4.24	1.45			2.4
S14	7.93	4.34	3.94;3.94			0.0
G15	8.11	3.97;3.89				0.6
I16	7.88	4.22	1.85	1.45;1.12	γCH_3:0.86;δCH_3:0.81	4.4
G17	8.40	4.20;4.02				5.8
P18		4.37	2.26;2.26	2.00;1.96	δδ′: 3.73;3.63	
E19	8.85	4.16	1.94;1.94	2.28;2.28		4.6
E20	8.09	4.18	1.90;1.88	2.18;2.18		1.8
H21	8.10	4.58	3.07;2.98		2H:8.35; 4H:7.13	4.0
F22	8.35	4.84	3.19;2.93		7.27(ave. ring)	6.0
P23		4.36	2.24; 2.24	1.95;1.93	δδ′: 3.69;3.55	
E24	8.56	4.23	2.08;1.97	2.32;2.32		6.1
NH_2	7.55;7.07					5.0;5.6
G17 CT	8.17	4.05;3.68				----
P18 CT		4.57	----;----	----;----		
E19 CT	8.71	4.27	2.01;2.01	2.27;2.27		----
E19 TC	8.85	4.22	2.03;2.03	2.28;2.28		4.6
E20 TC	8.11	4.25	2.00;1.93	2.24;2.24		1.8
H21 TC	8.22	4.63	3.22;3.06		2H:8.44; 4H:7.24	4.4
F22 TC	8.32	4.63	3.05;2.91			
P23 TC		3.78	1.86^{1};1.30^{2}	1.64^{1};1.61^{2}	$\delta\delta'$:3.52^{1}; 3.30^{2}	
E24 TC	8.80	4.13	2.01;2.01	2.30;2.30		5.9
NH_2 TC	7.78;7.06					5.7;6.0

[a] Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30ppm downfield from TMS.

[b] NH temperature coefficients are the result of a linear regression analysis of the chemical shifts at 5, 10, 15, 20 and 25 °C TOCSY spectra.

[c] The data were recorded at 500 MHz in CD_3OH/H_2O (60/40 v/v) at a concentration of 2mM. The pH was 5.5 in water, the temperature was 15 ±0.1 °C.

[d] Pairs of geminal protons x (= β, γ, δ, ε and α(G)) are not assigned stereospecifically, except those for *cis* Pro^{23}. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x′

[e] A doubling peaks due to R and S sulfoxide confurations were assigned as I and II according to the observed NOE patterns.

[f] Assignments may be interchanged and are ambiguous due to spectral overlap.

Table 2.4 ^{1}H NMR Resonance Assignments[a] and NH Temperature Coefficients[b] of the peptide NQ24[c].

Residue	NH	Hα [d]	Hβ [d]	Hγ [d]	Other H [d]	NH Temp. Coef. (-ppb/°C)
CH_3CO					CH_3CO: 2.01	
Q1	8.33	4.28	2.04;1.94	2.31;2.31	δNH_2: 7.68;6.86	4.7
K2	8.47	4.30	1.72;1.72	1.45;1.38	δδ′:1.57;1.57	4.3
					εε′:2.83; 2.74	
G3	8.89	3.86;4.18				7.1
L4	8.35	4.10	1.45; 1.45	1.37	δCH_3:0.88; 0.79	6.8
F5	8.34	4.50	3.15;2.97		7.23 (ave. ring)	8.9
D6	7.77	4.38	2.65;2.57			+0.7
F7	7.92	4.40	3.20;3.08		7.24 (ave. ring)	1.9
M8	8.09	4.23	2.57	2.44	εCH_3: 2.08	4.7
L9	7.97	4.20	1.73;1.73	1.58	δCH_3:0.87;0.83	3.4
E10	8.11	4.10	2.08;2.08	2.38;2.22		0.6
D11	8.26	4.52	2.70;2.70			5.0
E12	8.50	4.12	2.10; 2.10	2.35;2.27		1.6
A13	8.27	4.26	1.45			2.0
S14	7.93	4.34	3.95;3.95			0.0
G15	8.14	3.94;3.94				+0.8
I16	7.91	4.23	1.85	1.46;1.12	γCH_3:0.88; δCH_3:0.82	3.7
G17	8.49	4.22;4.01				5.3
P18		4.38	2.25;2.25	2.00;2.00	δδ′:3.74;3.63	
E19	8.97	4.14	1.95;1.95	2.26;2.26		4.7
E20	8.13	4.17	1.90;1.90	2.14;2.14		1.3
H21	7.98	4.51	2.94;2.90		2H:7.86; 4H:6.93	3.4
F22	8.21	4.87	3.17;2.91		7.27(ave. ring)	5.6
P23		4.38	2.23;2.23	1.95;1.95	δδ′: 3.63;3.56	
E24	8.60	4.20	1.95;1.95	2.06;2.27		5.7
NH_2	7.49;7.03					4.8;5.9
I16 CT	7.85	4.24	----;----	----;----	----;----	----
G17 CT	8.22	4.06;3.69				5.6
P18 CT		4.58	----;----	----;----	----;----	
E19 CT	8.80	4.23	1.93;1.93	2.25;2.25		----
E19 TC	8.96	4.20	1.95;1.95	2.26;2.26		4.7
E20 TC	8.14	4.22	1.96;1.96	2.30;2.30		2.0
H21 TC	----	----	----;----		2H:8.04; 4H:7.06	----
P23 TC	3.75	1.85 [1];1.29 [2]	1.64 [1];1.62 [2]		δδ′: 3.54[1];3.30[2]	
E24 TC	8.74	4.12	2.02;2.02	2.27;2.27		5.4
NH_2 TC	7.87;6.93					5.8;5.8

[a] Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30ppm downfield from TMS.

[b] NH temperature coefficients are the result of a linear regression analysis of the chemical shifts at 5, 10, 15, 20 and 25 °C TOCSY spectra.

[c] The data were recorded at 500 MHz in CD_3OH/H_2O (60/40 v/v) at a concentration of 1.7mM. The pH was 7.0 in 25mM phosphat buffer , the temperature was 15 ±0.1 °C.

[d] Pairs of geminal protons x (= β, γ, δ, ε and α(G)) are not assigned stereospecifically, except those for *cis* Pro[23]. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x′

Table 2.5 Distance constraints for the TC isomer of AD14 used for molecular modeling

Restrained Atoms	Dist(Å)	Restrained Atoms	Dist(Å)
CH_3CO-D1NH	2.6	A3αH-NH	2.5
D1αH-E2NH	2.4	I6αH-NH	2.5
E2αH-A3NH	2.3	P8αH-δ(3.70)	3.0
A3αH-S4NH	2.3	P8αH-δ(3.61)	3.2
S4αH-G5NH	2.3	E9αH-NH	2.3
G5ααH-I6NH	2.3	**D1αH-A3NH**	3.8
G7ααH-P8δδ	2.0	**E2αH-S4NH**	3.8
P8αH-E9NH	2.0	**A3αH-G5NH**	3.8
D1NH-E2NH	3.3	**S4αH-I6NH**	3.7
E2NH-A3NH	2.8	**G7α(4.01)H-E9NH**	3.7
A3NH-S4NH	2.5	**P8αH-E10NH**	3.1
S4NH-G5NH	2.5	**E2NH-S4NH**	4.6
G5NH-I6NH	2.4	**E9NH-E10NH**	2.5
G7NH-P8δ(3.70)	3.1	E10NH-H11NH	2.8
G7NH-P8δ(3.61)	3.3	P13δ(3.52)-E14NH	4.8
P8δ(3.70)-E9NH	2.8	E14NH-TNH(7.83)	3.5
P8δ(3.61)-E9NH	3.3	E10αH-H11NH	2.1
D1βH-E2NH	2.9	P13αH-F12β(3.06)	3.4
E2βH-A3NH	4.3	P13αH-F12β(2.92)	3.3
A3βH-S4NH	3.1	P13αH-E14NH	2.5
E10ββH-H11NH	3.5	E14α-NH	2.8
E2γ(2.08)-A3NH	3.5	E14αH-TNH (7.83)	2.8
E2γ(1.96) -A3NH	3.2	**F12αH-P13αH**	2.0
D1αH-NH	2.8	**F12αH-E14NH**	2.8
E2αH-NH	2.7		

Table 2.6 Distance constraints for the TT isomer of OQ24 used for molecular modeling

Restrained Atoms	Dist (Å)	Restrained Atoms	Dist (Å)
dNN(i,i+1)		E12-A13	2.3
Q1-K2	3.0	A13-S14	2.4
K2-G3	3.0	S14-G15	2.5
G3-L4	2.8	G15-I16	2.1
F5-D6	2.3	G17-P18	2.6
D6-F7	2.5	P18-E19	2.1
F7-M8	2.3	H21-F22	2.2
M8-L9	2.2	F22-P23	2.6
D11-E12	2.5	P23-E24	2.2
E12-A13	2.3	E24-TNH (7.56)	3.0
A13-S14	2.4	***dαβ(i,i+3)***	
S14-G15	2.3	G3(4.10)-D6	3.4
G15-I16	2.3	G3(3.87)-D6	4.1
I16-G17	2.5	L4-F7(3.18)	2.9
G17-P18 (3.72)	3.5	L4-F7(3.04)	2.9
G17-P18 (3.62)	3.6	F5-M8	2.5
P18(3.72)-E19	2.9	D6-L9	2.7
P18(3.62)-E19	3.5	F7-E10	2.7
E19-E20	2.4	M8-D11	2.6
H21-F22	2.5	D11-S14	3.2
E24-Ter NH	3.4	***Others***	
dαN(i,i+1)		K2α-D6β	2.9
Q1-K2	2.0	D6β-K2β	2.6
K2-G3	2.2	D6β-K2γ	2.9
G3(3.86)-L4	2.3	D6β-K2δ	3.0
F5-D6	2.4	G3NH-D6NH	3.8
M8-L9	2.3	L4α-F7NH	3.0
D11-E12	2.3	D11α-S14NH	3.2

Table 2.7 **Backbone torsion angles based on the minimized averaged structures obtained from molecular dynamics.**

Res.		AD14 solvated[a] A			NQ24 in vacuum B			NQ24 solvated C		
		ϕ	ψ	ω	ϕ	ψ	ω	ϕ	ψ	ω
1	Gln				-	173.8	-177.3	-	170.8	-169.8
2	Lys				153.5	-119.6	171.8	143.5	-87.8	173.1
3	Gly				93.1	-5.5	178.9	-114.2	-0.7	-179.1
4	Leu				-144.2	-56.0	177.6	-74.5	-34.4	178.5
5	Phe				-67.5	-54.5	173.8	-59.6	-51.9	176.0
6	Asp				-63.4	-30.4	173.9	-65.7	-7.2	173.7
7	Phe				-67.3	-8.9	172.4	-79.2	-16.0	171.9
8	Met				-91.2	-15.2	172.3	-74.9	-41.8	169.9
9	Leu				-84.3	-44.0	169.7	-56.5	-53.1	175.0
10	Glu				-65.8	-4.6	175.0	-69.0	3.6	172.8
11	Asp	-	-5.6	+174.2	-87.4	-45.5	169.3	-99.3	-47.1	170.3
12	Glu	-99.6	-6.6	-175.7	-61.0	-7.6	173.0	-57.6	-36.5	170.7
13	Ala	-104.5	-67.0	174.8	-117.0	9.4	-175.2	-50.4	-19.6	175.8
14	Ser	-122.7	29.3	-174.3	-49.0	-33.2	177.9	-61.4	-18.8	172.6
15	Gly	-61.6	-9.4	173.2	-58.4	-13.7	170.4	-62.1	-13.4	171.0
16	Ile	-72.3	57.1	172.1	-131.5	12.7	-168.4	-130.6	11.0	-168.2
17	Gly	114.5	-74.1	-167.2	-82.8	-67.1	178.5	-81.8	-67.5	-179.3
18	Pro	-79.6	51.2	162.0	-87.7	-119.2	-175.8	-85.3	-119.9	-176.7
19	Glu	-121.6	-22.6	179.3	-156.4	-2.5	179.0	-155.2	-5.7	-179.8
20	Glu	-108.2	122.1	-177.3	-70.7	174.6	-167.9	68.0	174.0	-167.9
21	His	146.8	39.6	173.8	-62.5	-28.2	-172.8	-61.7	-28.9	-172.5
22	Phe	-142.7	78.6	-31.6	79.6	106.2	-169.7	79.9	107.4	-169.7
23	Pro	-67.9	-3.9	173.1	-74.8	-5.6	-177.2	-74.9	-5.4	-177.6
24	Glu	-132.7	-	-	-115.3	-	-	-113.6	-	-

[a]The residue numbers are from 1 to 14.

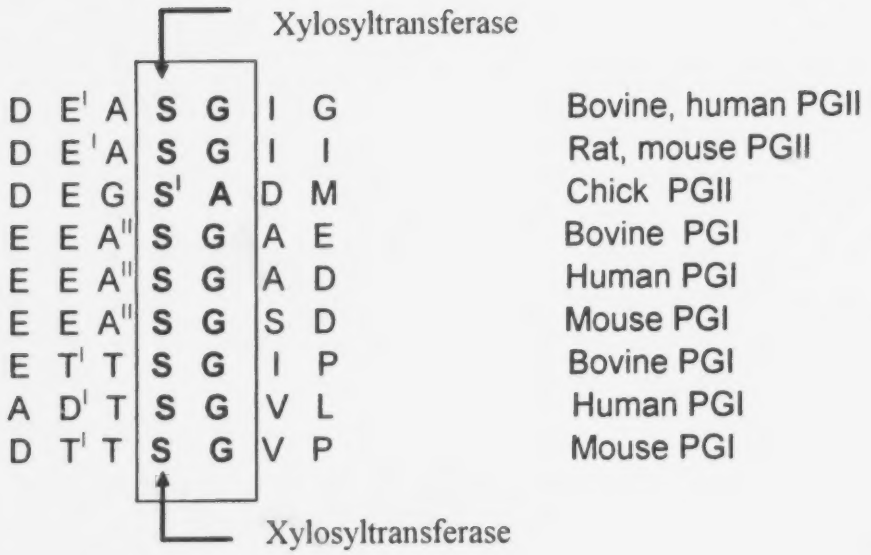

Xylosyltransferase
D E[I] A S G I G Bovine, human PGII
D E[I] A S G I I Rat, mouse PGII
D E G S[I] A D M Chick PGII
E E A[II] S G A E Bovine PGI
E E A[II] S G A D Human PGI
E E A[II] S G S D Mouse PGI
E T[I] T S G I P Bovine PGI
A D[I] T S G V L Human PGI
D T[I] T S G V P Mouse PGI
Xylosyltransferase

Figure 2.1. Glycosylated sequences in small DS-PGs, superscripts are β-turn predictions (7)

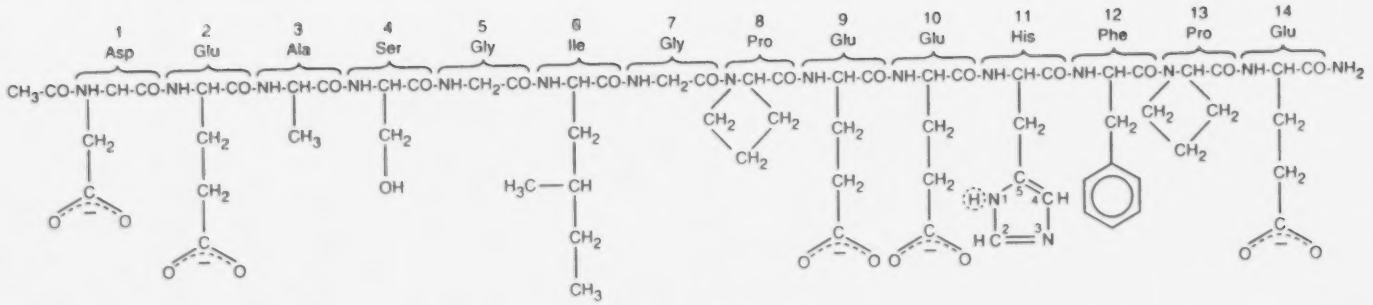


Figure 2.2 The structure of the 14-mer peptide, AD14

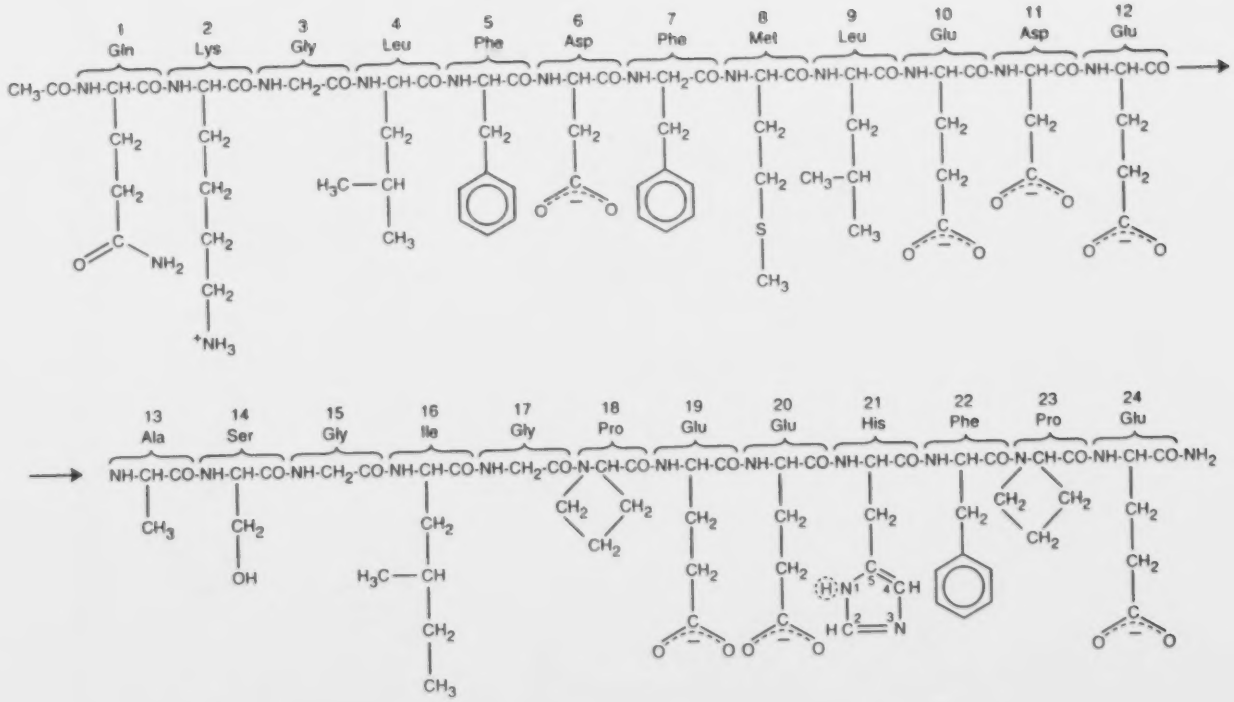


Figure 2.3 The structure of the 24-mer peptide, NQ24

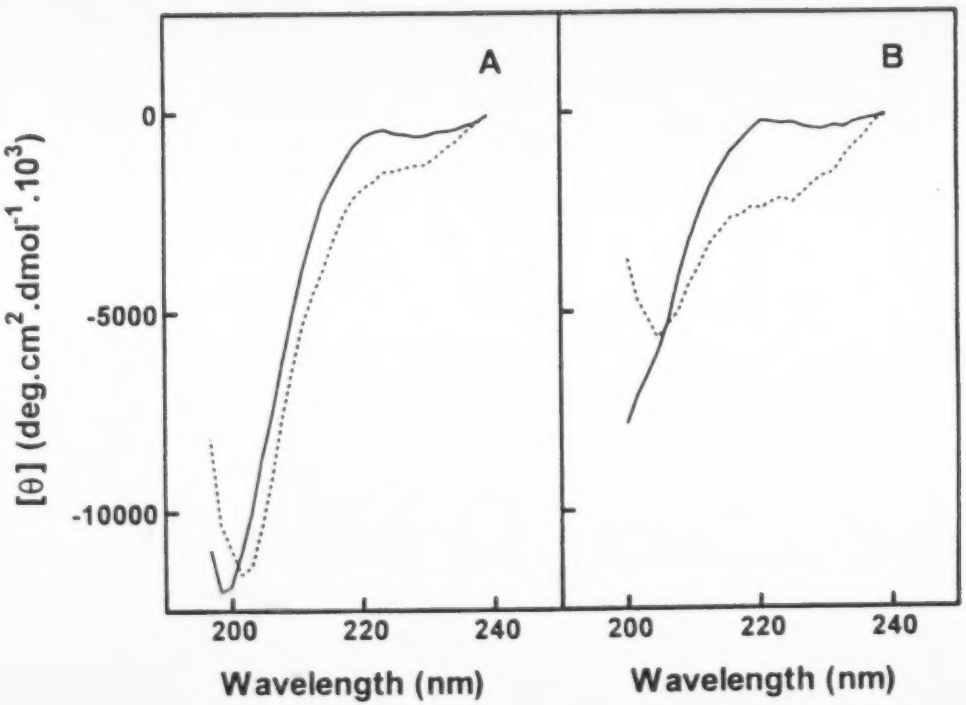


Figure 2.4 Circular dichroism spectra of D14 (A) and NQ24 (B) in 0,05M sodium phosphate buffer (----) pH 5.5 or in 60% methanol/40% sodium phosphate buffer(- - -). See text for other details.

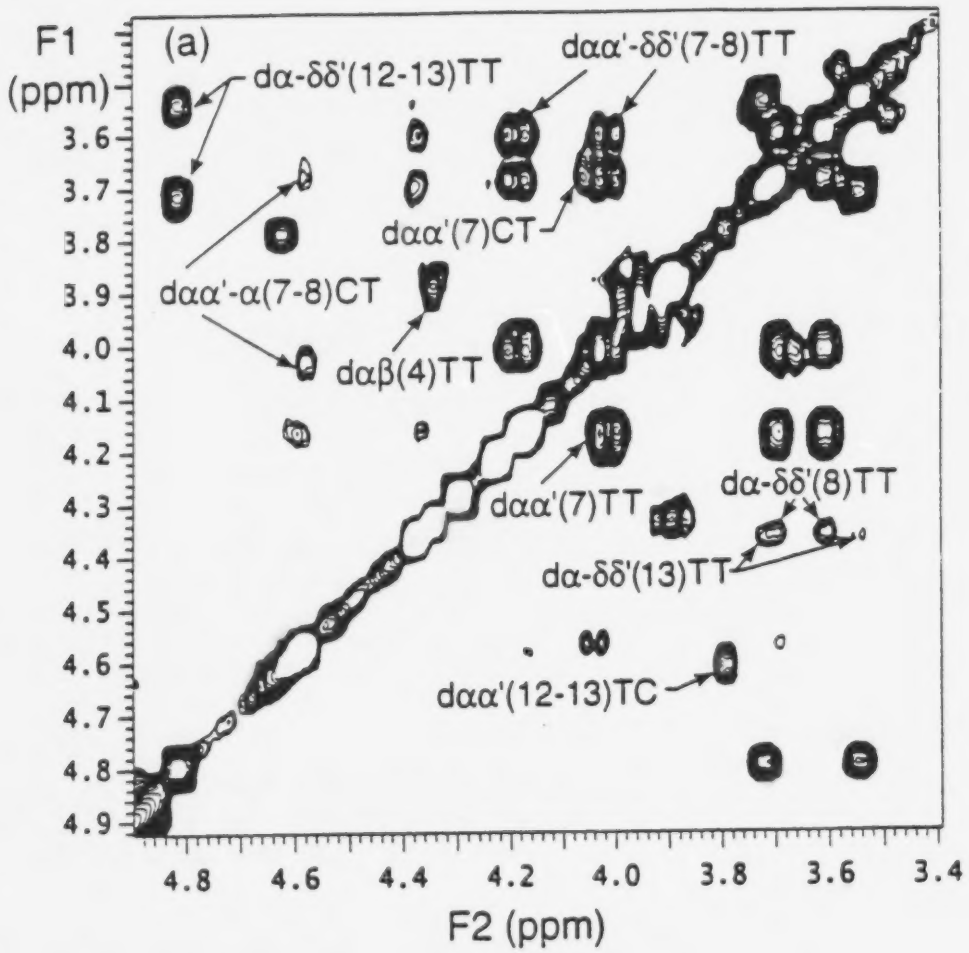


Figure 2.5 The αβ region of the 500 MHz NOESY NMR spectrum of the peptide AD14. The concentration was 2.3mM, at pH 5.5, in CD_3OH/H_2O (60/40 v/v) solvent, the temperature was 10±0.1°C and mixing time was 250 ms. The labeling of the protons is as follows: The protons giving rise to the cross-peak are defined first, followed by the amino acid residue number and ending with the *cis* or *trans* designation for Pro^8 and Pro^{13}. This convention is used in Figures 2.10; 2.12 and 2.14.

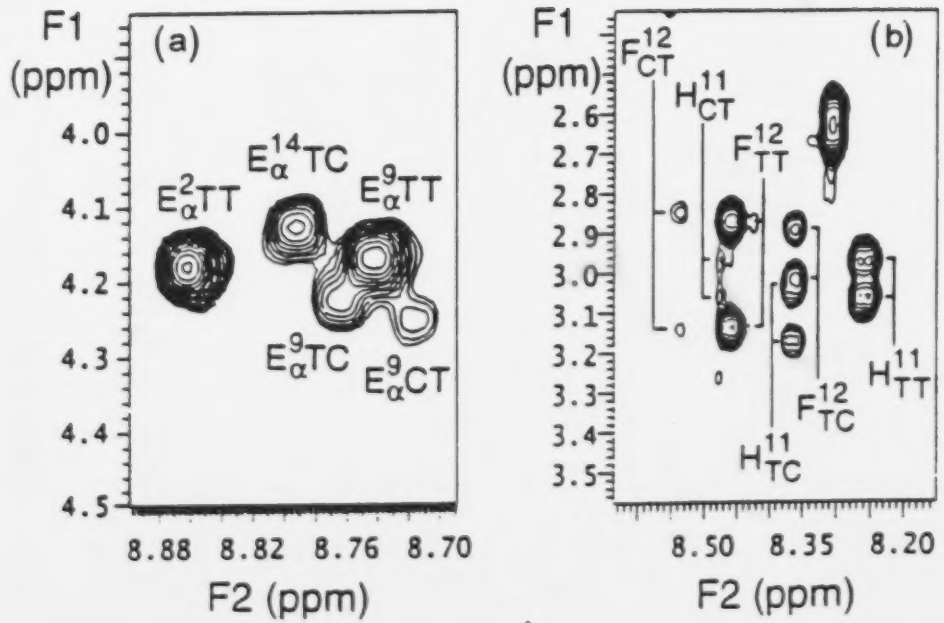


Figure 2.6 The αH-NH (a) and βH-NH (b) region of the TOCSY spectra of peptide AD14.

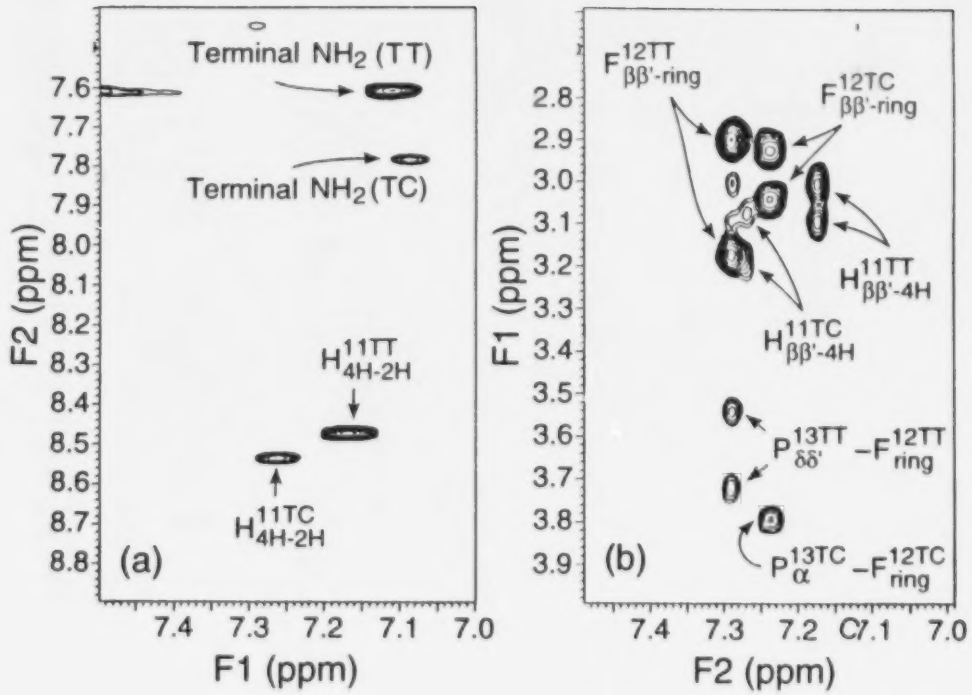


Figure 2.7 The NH-NH region of the TOCSY spectrum (a) and the βH-NH region of the NOESY spectrum (b) of the peptide AD14.

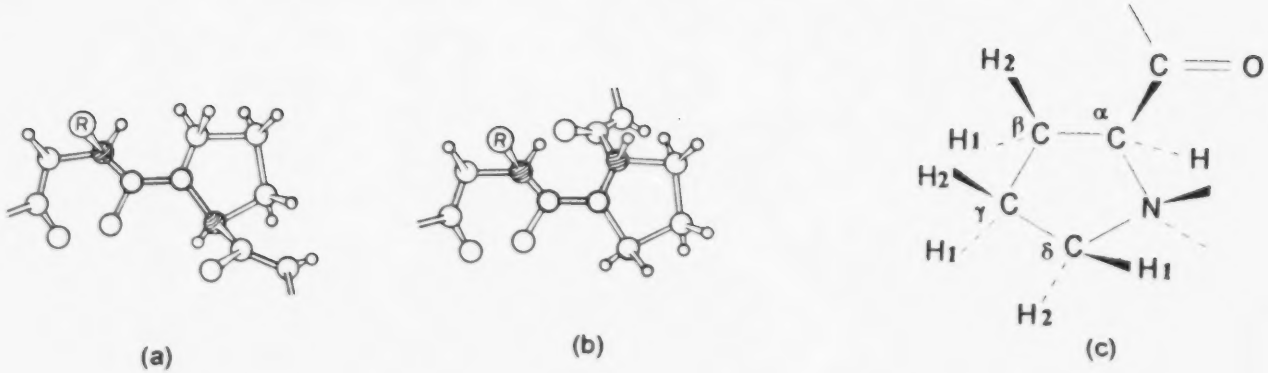


Figure 2.8 Schematic diagram of *trans* (a), *cis* (b) proline(24) and the assignment (25) of its protons (c)

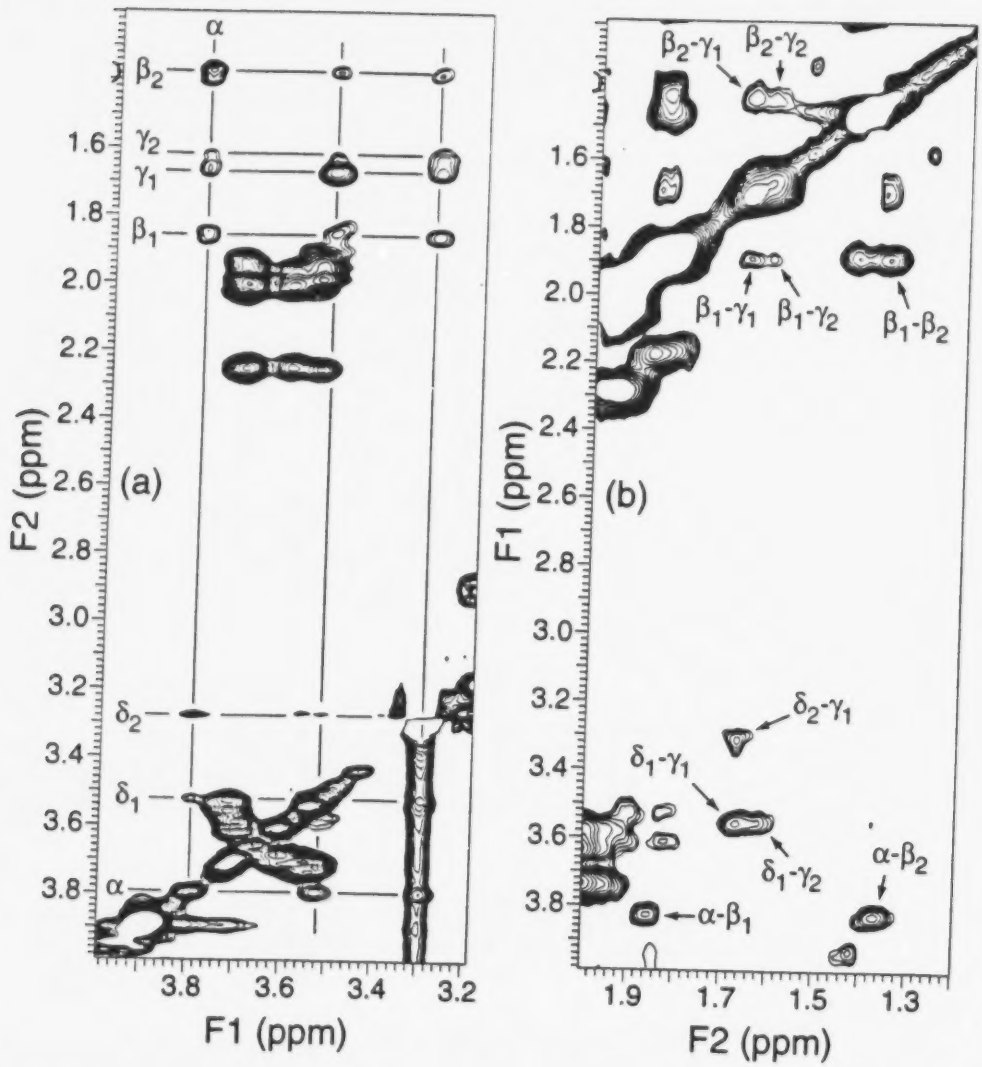


Figure 2.9 Sections of TOCSY (a) and NOESY (b) spectra of *cis* Pro[13] of peptide AD14.

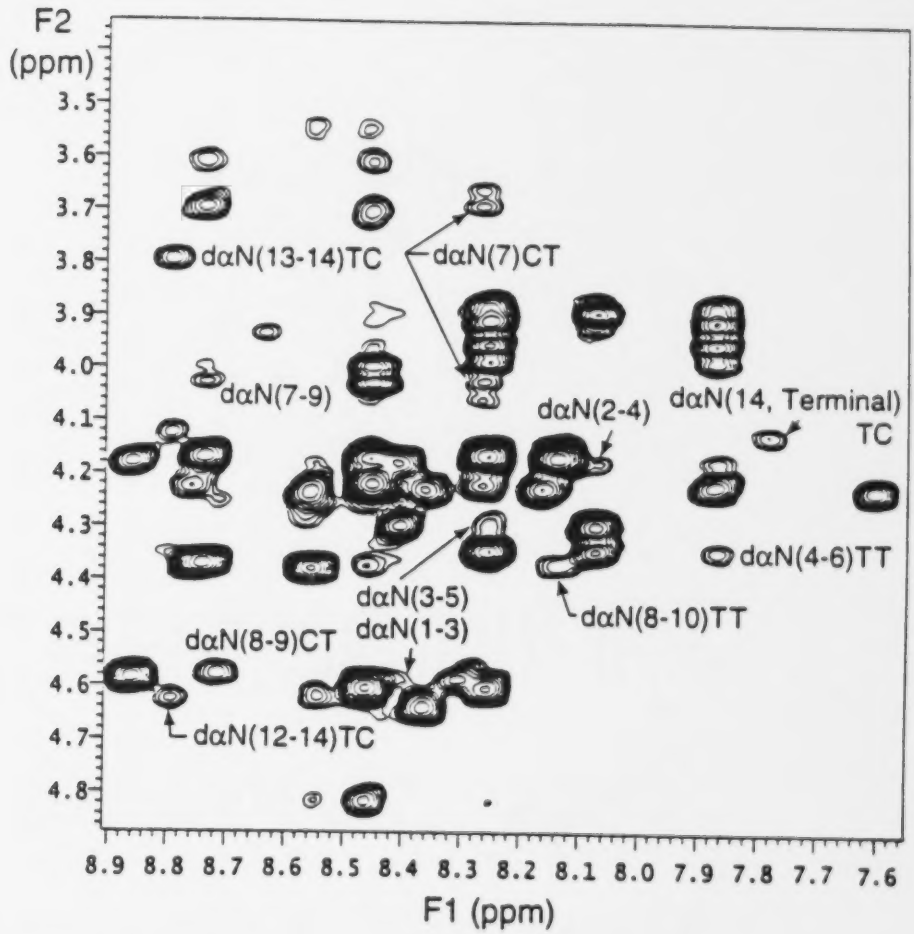


Figure 2.10 The αH-NH region of the NOESY spectra of the peptide AD14

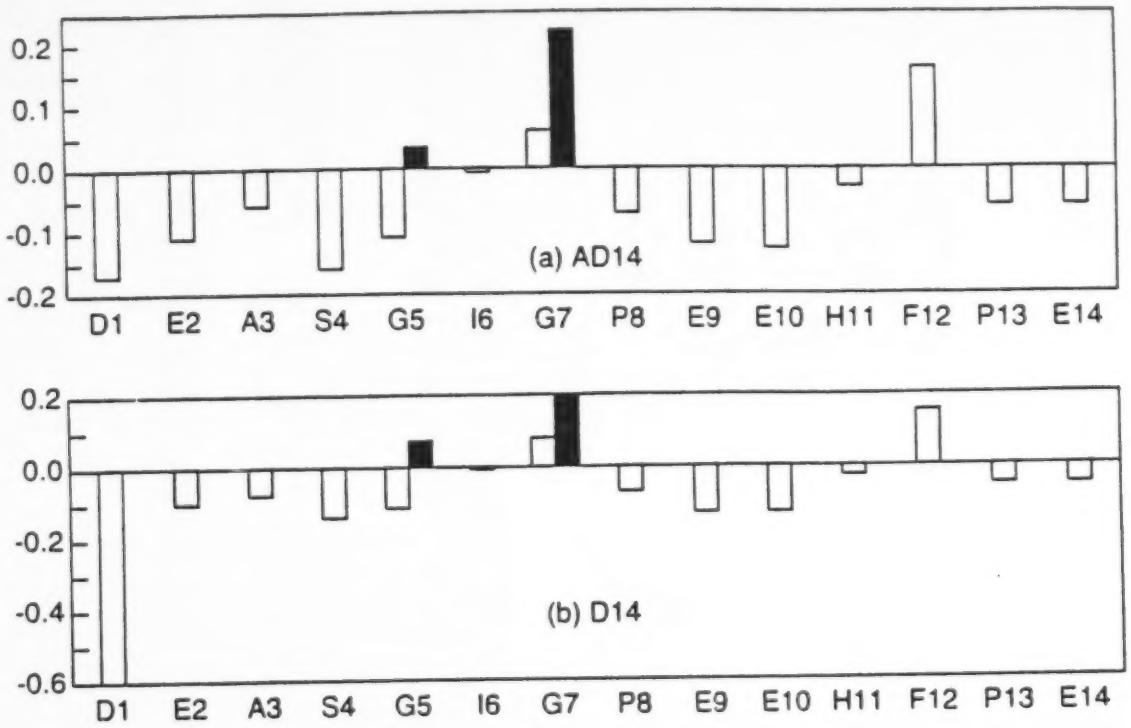


Figure 2.11 Net structural shifts of the α protons for the peptides (a) AD14 and (b) D14

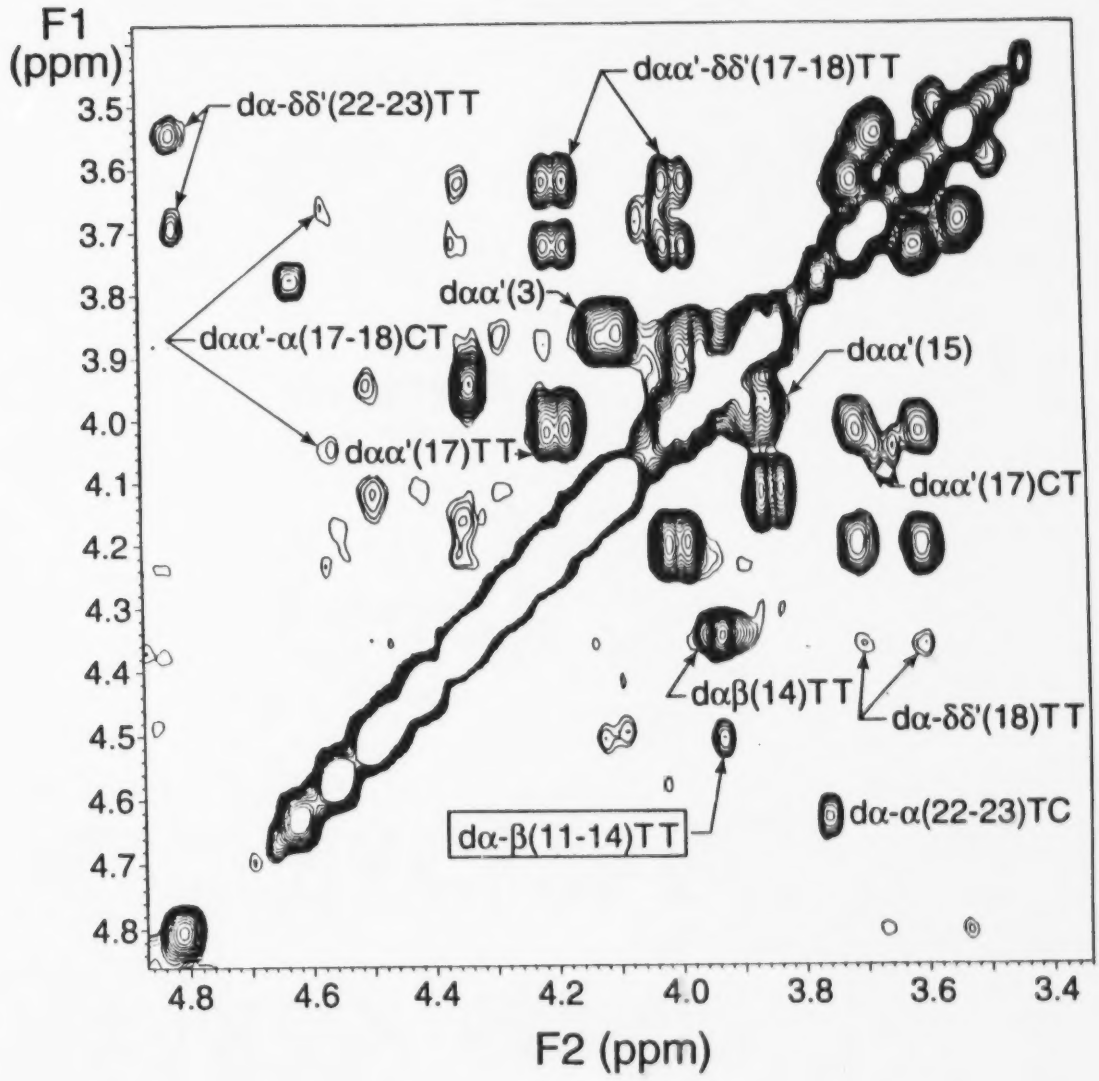


Figure 2.12 The αH-βH region of the NOESY spectra of peptide OQ24

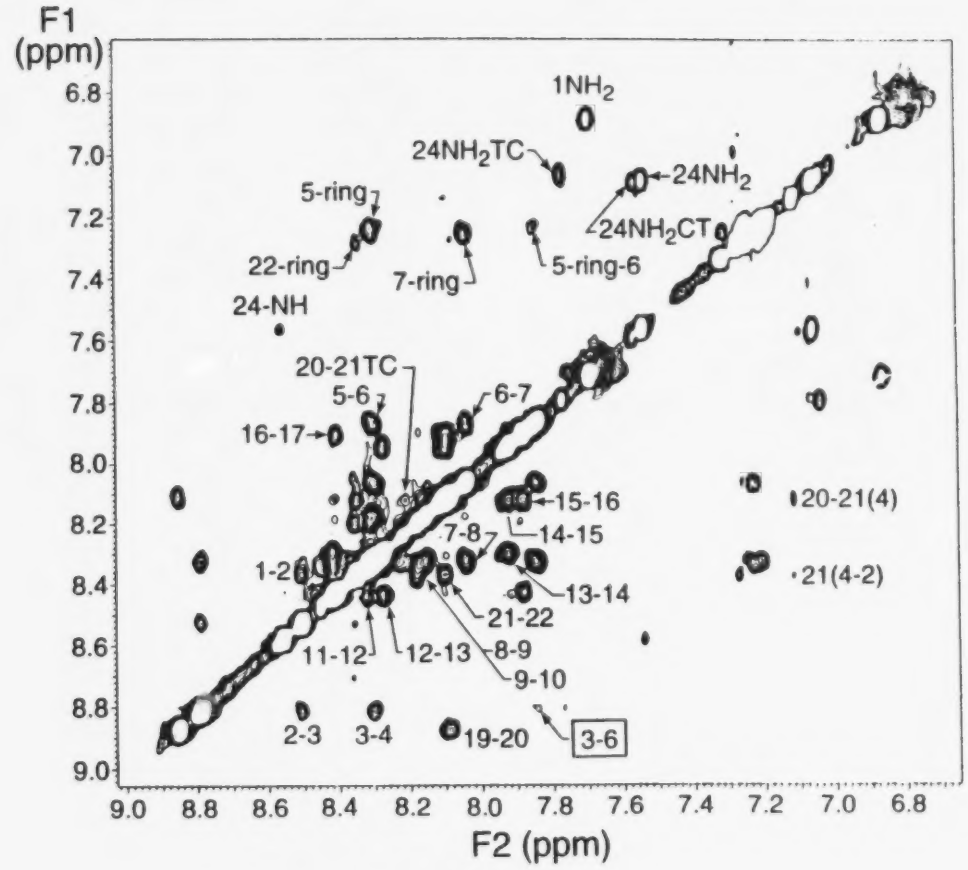


Figure 2.13 The NH-NH region of the NOESY spectra of peptide OQ24.

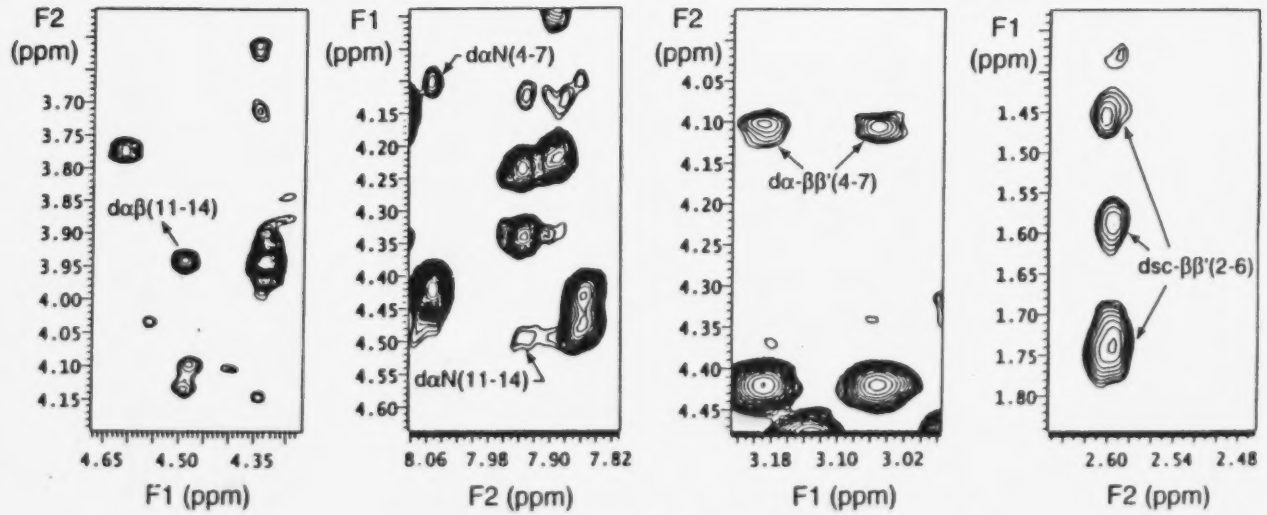


Figure 2.14 Sections of the NOESY spectra of peptide OQ24

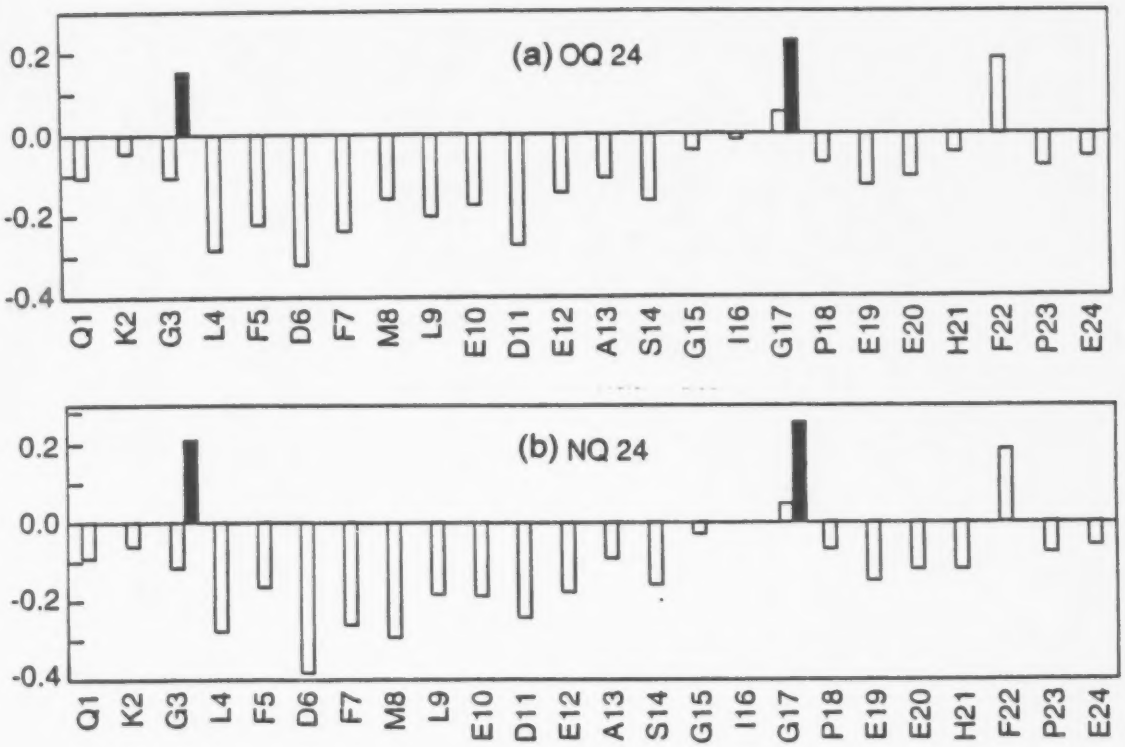


Figure 2.15 Net structural shifts of the α protons for the peptides (a) OQ24 and (b) NQ24

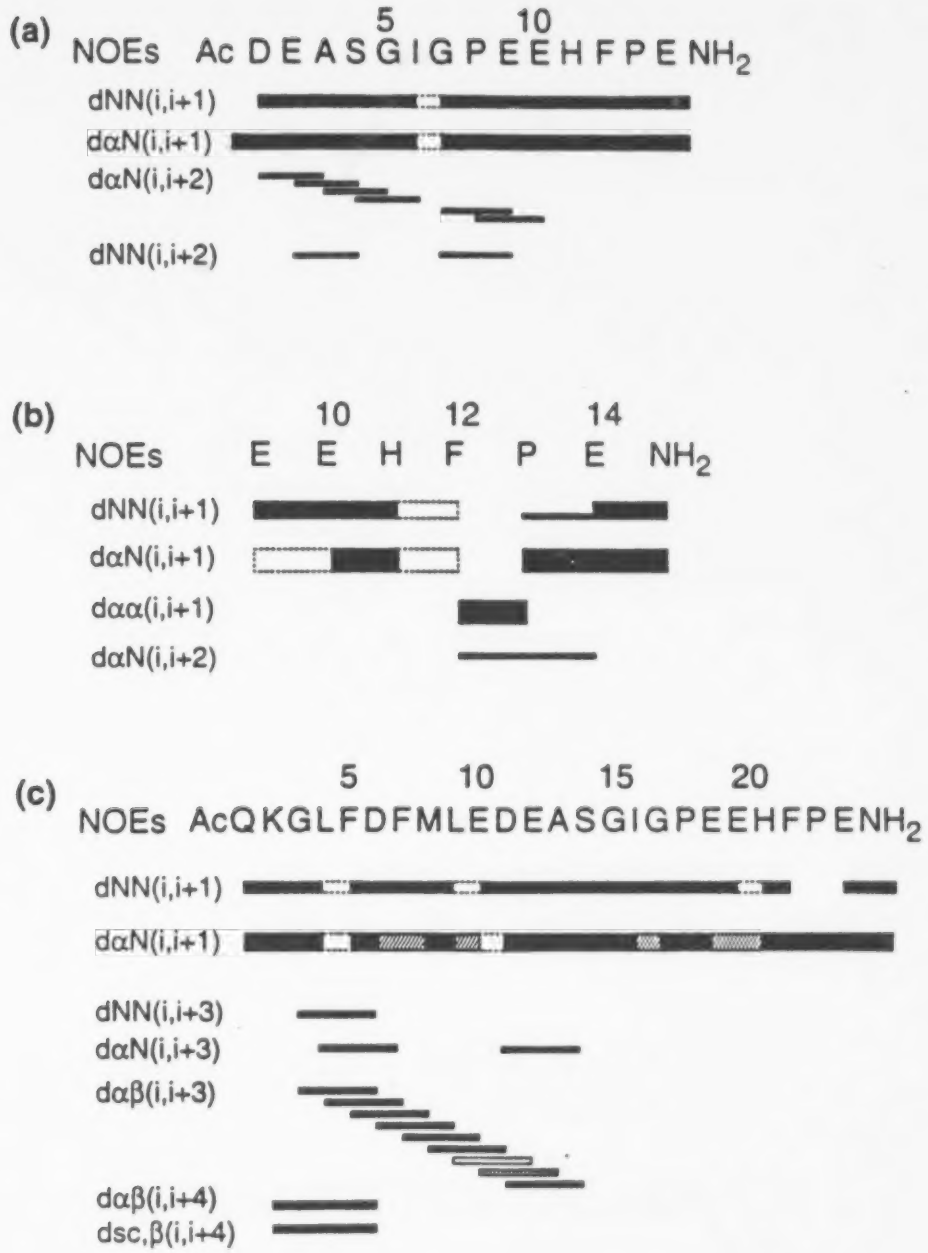


Figure 1.16 Survey of NOE effects recorded for (a) AD14 (b) part of AD14 TC isomer, (c) OQ24. For the proline, δ-protons are used instead of the NH-proton. Solid (■) indicate unambiguous NOE's Shadowed lines (▤) indicate NOE's that are ambiguous due spectra overlap, and blank space (□) indicate an NOE unobservable due to the overlap. The thickness of the bars indicates the strength of the NOE.

Figure 2.17 The last 20 conformations from the 100 ps dynamics run for peptide AD14 with *cis* proline in position 13

Figure 2.18 Averaged minimized structure for the 24-mer with all bonds *trans* modeled without solvent

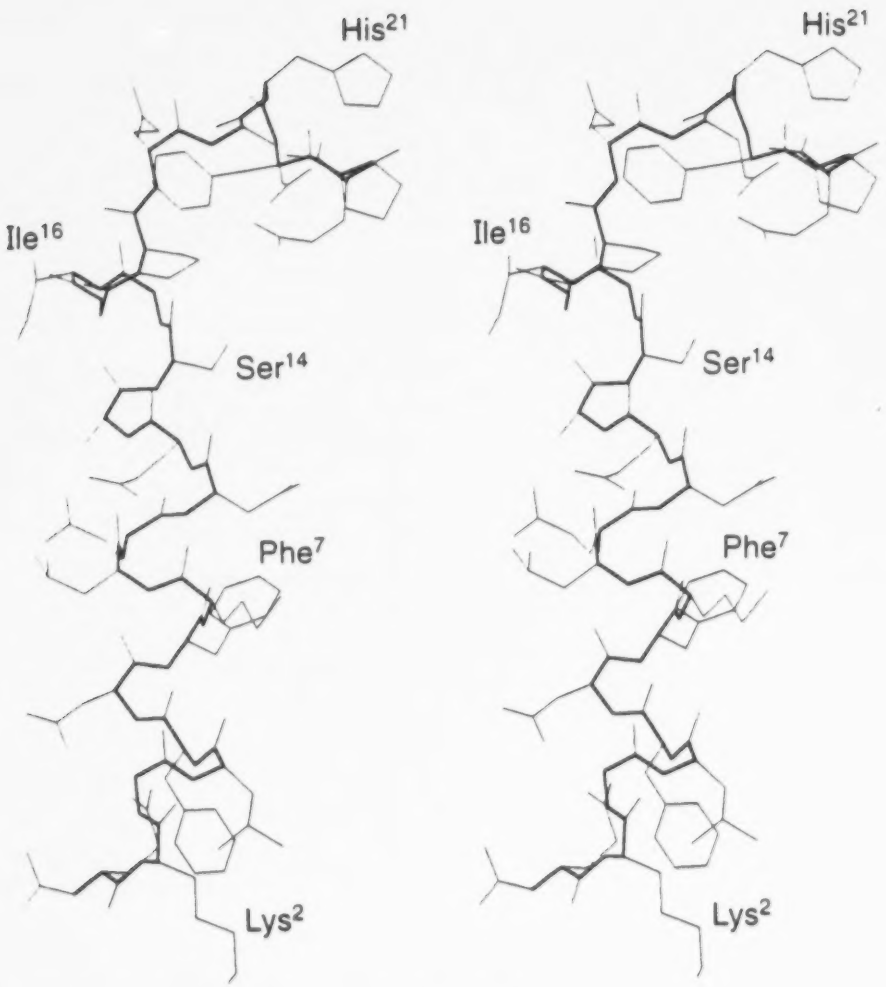


Figure 2.19 The stereo view of the average minimized structure of the 24-mer after modeling with the shell of solvent